EUROPEAN PATENT APPLICATION

(54) Method for speech coding and voice-coder.

(57) A method for speech coding and a voice-coder for coding speech signals divided into frames spaced with a constant interval are disclosed. An adaptive codebook storing excitation signal determined in advance and a plurality of excitation codebooks for multi-stage vector quantization are provided. Each frame is divided into subframes. For each subframe, a candidate of a first predetermined number of adaptive codevectors is selected, and then candidates of each predetermined number of excitation codevectors are selected from each excitation codebook, respectively, by using the candidate of the adaptive codevector. Finally, a combination of the adaptive codevector and each of the excitation codevector is selected from the candidates of the adaptive codevector and each of the sound codevectors.
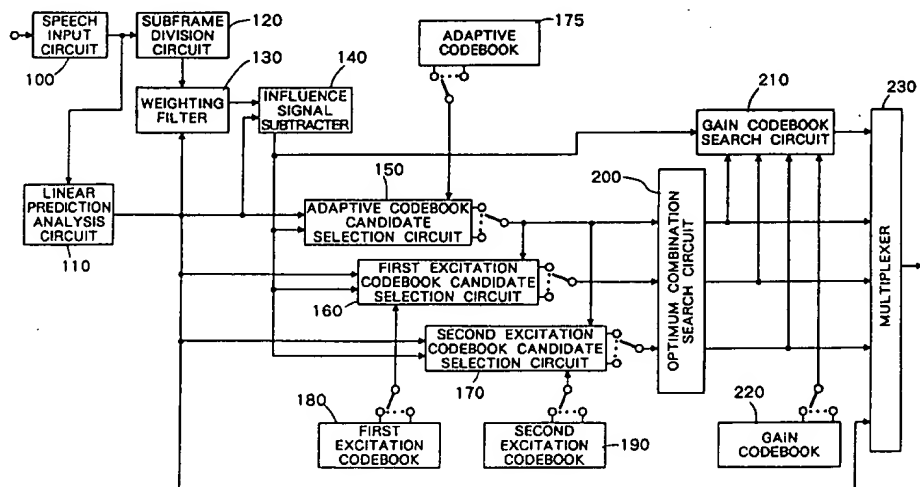
EP 0 545 386 A2

FIG. 1

The present invention relates to a method for speech coding and to a voice-coder, particularly to a method for speech coding and to a voice-coder which can achieve high coding quality with a relatively small operation at bit rates not greater than 8 kbit/s.

As a speech coding system to be applied to vector quantization of excitation signals at low bit rates by using a excitation codebook comprising random numbers, a CELP system, described in a paper (hereinafter referred to as literature 1) titled "CODE-EXCITED LINEAR PREDICTION (CELP): HIGH-QUALITY SPEECH AT VERY LOW BIT RATES" (Proc. ICASSP, pp. 937-940, 1985) by Manfred R. Shroeder and Bishnu S. Atal is known. There is also a CELP system using an adaptive codebook described in a paper (hereinafter referred to as literature 2) titled "IMPROVED SPEECH QUALITY AND EFFICIENT VECTOR QUANTIZA-TION IN SELP" (Proc. ICASSP, pp. 155-158, 1988) by W. B. Kleijin, D. J. Krasinski and R. H. Ketchum. The CELP system using the adaptive codebook receives speech signals divided into frames spaced with a constant interval. The CELP utilizes a linear predictive analyzer for obtaining spectral parameters of input speech signals, the adaptive codebook having excitation signals determined in the past, the excitation codebook comprising random numbers to be used for vector quantization of the excitation signals of said input speech signals. The CELP selects an adaptive codevector by using the input speech signal and the synthesized signal of the adaptive codevector for every subframe made by equally dividing the frame. Subsequently, CELP performs selection of excitation codevectors by using the input signals, the syn-thesized signal of the selected adaptive codevector and said excitation codevector.

However, the CELP systems have the following disadvantage, in that a quite large operation is required for searching the excitation codebook. Moreover since the adaptive codebook is determined independently of the excitation codebook, it is impossible to get a high SN (signal to noise) ratio. Further in the above CELP system, although the adaptive codebook and the excitation codebook are each searched by using gains not quantized, it becomes possible to obtain a higher SN ratio when the adaptive codebook and the excitation codebook are searched for all the quantization value of gains. Furthermore, it is impossible to obtain sufficiently good speech quality with low bit rates such as 8 kbit/s or less because of the too small size of the excitation codebook.

An object of the present invention is to provide a method for speech coding which can solve the above problem of the conventional method and achieve high quality speech by a relatively small operation even at the low bit rates such as less than 8 kbit/s.

Another object of the present invention is to provide a voice-coder which can solve the above problem of the conventional method and achieve high quality speech by a relatively small operation even at low bit rates such as less than 8 kbit/s.

The object of the present invention can be achieved by a method for speech coding for coding speech signals divided into frames spaced with a constant interval, wherein an adaptive codebook storing excitation signals determined in the past and a plurality of excitation codebooks for multi-stage vector quantization of an excitation signal of the input speech signal are prepared; a spectral parameter of said input speech signal is obtained; said frame is divided into subframes; a candidate of a first fixed number of adaptive codevectors is selected for every said subframe from said adaptive codebook by using said input speech signal and said spectral parameter; candidates of a second fixed number of excitation codevectors are selected for every said subframe from said excitation codebooks, respectively, by using said input speech signal, said spectral parameter and the candidate of said adaptive codevector; and a combination of the adaptive vector and each of the excitation codevectors forming an excitation signal of said subframe is selected from the candidates of said adaptive codevector and each of said excitation codevector by using said input speech signal and said spectral parameter.

Another object of the present invention is achieved by a voice-coder for coding speech signals divided into frames spaced with a constant interval, comprising: linear prediction analysis means for outputting spectral parameters of input speech signals; an adaptive codebook for storing excitation signals determined in the past; a plurality of excitation codebooks provided for multi-stage vector quantization of the excitation signal of said input speech signals; wherein, in case of searching for a combination of the adaptive codevector and each of the excitation codevectors for every subframe prepared by further division of said frame, from said adaptive codebook and each of said excitation codebooks, respectively, said combination of the adaptive codevector and each of the excitation codevectors forming an excitation signal of said subframe; a candidate of a first predetermined number of adaptive codevectors is selected from said adaptive codebook by using said input speech signal and said spectral parameter; candidates of each predetermined number of excitation codevectors are selected from a plurality of said excitation codebooks respectively by using said input speech signal, said spectral parameter and the candidate of said adaptive codevector; and a candidate of the adaptive codevector and each of excitation codevectors forming the excitation signal of said subframe is selected from the candidate of said adaptive codevector and each of

2

said excitation codevectors by using said input speech signal and said spectral parameter.

Another object of the present invention is also achieved by a voice-coder for coding speech signals divided into frames spaced with a constant interval, comprising: linear prediction analysis means for outputting spectral parameters of input speech signals; an adaptive codebook storing excitation signals determined in the past; a plurality of excitation codebooks provided for multi-stage vector quantization of an excitation signal of said input speech signals; subframe division means for generating subframe signals by dividing said frame into subframes; first selection means for selecting a candidate of a first fixed number of adaptive codevectors from said adaptive codebook in accordance with said subframe signal and said spectral parameter; second selection means provided for every said excitation codebook for selecting the candidate of the excitation codevectors of the number predetermined for every excitation codebook, from the corresponding excitation codebook in accordance with said subframe signal, said spectral parameter and the candidate of said adaptive codevector; and means for searching the candidate of the adaptive vector and each of the excitation codevectors which forms the excitation signal of said subframe, from the candidate of said adaptive codevector and the candidate of each of said excitation codevectors in accordance with said input aural signal and said spectral parameter.

The above and other objects, features and advantages of the present invention will be apparent from the following description referring to the accompanying drawings which illustrate examples of preferred embodiments of the present invention.

BRIEF DESCRIPTION OF THE DRAWINGS:

Fig. 1 is a block diagram showing the structure of a voice-coder of a first embodiment of the present invention.

Fig. 2 is a block diagram showing the structure of a voice-coder of a second embodiment of the present invention.

Fig. 3 is a block diagram showing the structure of a voice-coder of a third embodiment of the present invention.

Fig. 4 is a block diagram showing the structure of a voice-coder of a fourth embodiment of the present invention.

DESCRIPTION OF THE PREFERRED EMBODIMENTS:

A first preferable embodiment of the present invention will be described with reference to Fig. 1. In the voice-coder shown in Fig.1, there are provided adaptive codebook 175, gain codebook 220 and two kinds of excitation codebooks 180, 190.

Speech input circuit 100 is provided for receiving speech signals divided into frames spaced with a constant interval. Subframe division circuit 120 and linear prediction analysis circuit 110 are provided on the output side of speech input circuit 100. Subframe division circuit 120 outputs subframes by equally dividing the frame, and linear prediction analysis circuit 110 performs linear prediction analyses of speech signals composing frames for obtaining spectral parameters of the speech signals. On the output side of subframe division circuit 120, weighting filter 130 is provided for performing perceptual weighting by receiving subframes and spectral parameters. On the output side of weighting filter 130, influence signal subtracter 140 is provided for subtracting weighted influence signal from the preceding subframe and outputting the results thereof.

Adaptive codebook 175 stores excitation signals decided in the past as adaptive codevectors. Corresponding to adaptive codebook 175, adaptive codebook candidate selection circuit 150 is provided for selecting the previously fixed number of adaptive codevectors and for outputting thereof as candidates of adaptive codevectors. Adaptive codebook candidate selection circuit 150 performs selection of the candidate according to the spectral parameter and the output signal of influence signal subtracter 140.

First and second excitation codebooks 180, 190 operate for multi-stage vector quantization of the excitation signal, and store the first and second excitation codevectors, respectively. Corresponding to first and second excitation codebooks 180, 190, candidate selection circuits 160, 170 for the first and second excitation codebooks are provided respectively. Candidate selection circuits 160, 170 select the previously fixed number of excitation codevectors from corresponding respective excitation codebooks 180, 190 and output thereof as the candidates of the excitation codevectors. Spectral parameters, output signals of the influence signal subtracter and candidates of adaptive codevectors are inputted into each of candidate selection circuits 160, 170 for the excitation codebook.

Optimum combination search circuit 200 is provided for candidates selected by candidate selection circuits 150, 160, 170 for the corresponding codebooks in order to search the optimum combination of candidates. Further, gain codebook search circuit 210 and multiplexer 230 are provided. Optimum combination search circuit 200 is structured so as to output to multiplexer 230 the delay (to the adaptive codevector) or index (to each excitation codevector) with reference to each of the respective optimum adaptive codevectors, to the first and second excitation codevectors according to the results of the search, and to the output weighted synthetic signals of the above vectors to gain codebook search circuit 210, respectively. Gain codebook search circuit 210 searches for the optimum gain codevector from gain codebook 220 which stores gain codevectors, and outputs the index of thus searched optimum gain codevector. Multiplexer 230 is structured so as to receive delay or indices from optimum combination search circuit 200 or gain codebook search circuit 210, and output codes which correspond to input speech signals according to delay or indices.

Next, description will be made with reference to selection or search algorithm of each candidate selection circuit 150, 160, 170 or optimum combination search circuit 200 of the present embodiment. Under these algorithms, the excitation signal is processed by two-stage vector quantization by using two kinds of excitation codebooks 180, 190.

First, in adaptive codebook candidate selection circuit 150, the predetermined number L_0 of the adaptive codevectors is selected, in order, from the one with smaller error E_0 expressed by equation (1):

$$E_0 = \| z - \beta_0 sa_d \|^2 \quad (1)$$

where, z is a signal obtained by subtracting an influence signal from a perceptually weighted input signal, sa_d a perceptually weighted synthetic signal of adaptive codevector a_d with delay d, β_0 a sequential optimum gain of an adaptive codevector, $\| \ \|$ Euclid norm. The sequential optimum gain β_0 of the adaptive codevector is given by:

$$\beta_0 = \frac{<z, \ sa_d>}{<sa_d, \ sa_d>} \quad (2)$$

By substituting above equation (2) into equation (1), the next equation is obtained.

$$E_0 = \| z \|^2 - \frac{<z, \ sa_d>^2}{<sa_d, \ sa_d>} \quad (3)$$

where, $<, >$ represents an inner product.

In candidate selection circuit 160 for the first excitation codebook, candidates of predetermined number L_1 of the first excitation codevectors are selected for each L_0 piece of the adaptive codevectors selected by candidate selection circuit 150 for the adaptive codebook, in order, from the one with smaller error E_1 expressed by equation (4):

$$E_1 = \| za - \gamma_0 se_i^1 \|^2 \quad (4)$$

where se_i^1 is a perceptually weighted synthesized signal of first excitation codevector e_i^1 with index i, γ_0 a sequential optimum gain of the first excitation codevector, and $za = z - \beta_0 sa_d$.

Therefore:

$$\gamma_0 = \frac{<za, \ se_i^1>}{<se_i^1, \ se_i^1>} \quad (5)$$

4

By substituting above equation (5) into equation (4), equation (6) below is obtained:

$$E_1 = \|z\|^2 - \frac{<z, \; se_i^1>^2}{<se_i^1, \; se_i^1>} \qquad (6)$$

In the same way as described above, in candidate selection circuit 170 for the second excitation codebook, the candidates of predetermined number L_2 of the second excitation codevectors are selected for each L_0 piece of adaptive codevectors selected by candidate selection circuit 150 for the adaptive codebook, in order, from the one with smaller error E_2 expressed by the next equation:

$$E_2 = \|za - \delta_0 se_j^2\|^2 \qquad (7)$$

where se_j^2 is a perceptually weighted synthesized signal of second excitation codevector e_j^2 with index j, and δ_0 a sequential optimum gain of the second excitation codevector. Therefore:

$$\delta_0 = \frac{<za, \; se_j^2>}{<se_j^2, \; se_j^2>} \qquad (8)$$

By substituting equation (8) into equation (7), following equation (9) is obtained.

$$E_2 = \|za\|^2 - \frac{<z, \; se_j^2>^2}{<se_j^2, \; se_j^2>} \qquad (9)$$

In optimum combination search circuit 200, error E is calculated by the following equation for all the combinations of candidates of the selected adaptive codevectors, and the first and second excitation codevectors, and then the combination of the candidates with minimum E is searched.

$$E = \|z - \beta sa_d - \gamma se_i^1 - \delta se_j^2\|^2 \qquad (10)$$

where β, γ, δ are simultaneous optimum gains of an adaptive codevector, the first and second excitation codevectors, respectively. Therefore:

$$\begin{pmatrix} \beta \\ \gamma \\ \delta \end{pmatrix} = R^{-1} \begin{pmatrix} <z, \; sa_d> \\ <z, \; se_i^1> \\ <z, \; sa_j^2> \end{pmatrix} \qquad (11)$$

However, R is to satisfy the following equation:

$$R = \begin{pmatrix} <sa_d, sa_d> & <sa_d, se_i^1> & <sa_d, se_j^2> \\ <se_i^1, sa_d> & <se_i^1, se_i^1> & <se_i^1, se_j^2> \\ <se_j^2, sa_d> & <se_j^2, se_i^1> & <se_j^2, se_j^2> \end{pmatrix}$$

---------- (12)

By substituting equation (11) into equation (10), thus obtained,

$$R = \|z\|^2 - (<z, sa_d> \ <z, se_i^1> \ <z, se_j^2>) \times R^{-1} \begin{pmatrix} <z, sa_d> \\ <z, se_i^1> \\ <z, se_j^2> \end{pmatrix}$$

---------(13)

When above error E is calculated, it is acceptable to assign a particular limitation to simultaneous optimum gains γ, δ of each excitation codevector. For example, with the limitation that γ and δ are equal, error E is given by,

$$E = \|z\|^2 - (<z, sa_d> \ <z, se_i^1 + se_j^2>) \times R^{-1} \begin{pmatrix} <z, sa_d> \\ <z, se_i^1 + se_j^2> \end{pmatrix}$$

---------(14)

where,

$$R = \begin{pmatrix} <sa_d, sa_d> & <sa_d, se_i^1 + se_j^2> \\ <se_i^1 + se_j^2, sa_d> & <se_i^1 + se_j^2, se_i^1 + se_j^2> \end{pmatrix}$$

---------(15)

Next, description will be made with reference to operation of the voice-coder of the present embodiment.

Speech input circuit 100 receives speech signals divided into each frame (e.g., 40 ms in width), which signals are outputted to linear prediction analysis circuit 110 and subframe division circuit 120. In linear prediction analysis circuit 110, linear prediction analysis of the inputted speech signal is performed for calculating the spectral parameter. This spectral parameter is outputted to weighting filter 130, to influence signal subtracter 140, to candidate selection circuit 150 for the adaptive codebook, to candidate selection circuit 160 for the first excitation codebook, to candidate selection circuit 170 for the second excitation codebook, and to multiplexer 230. Separately, a frame is divided into subframes (e.g., 8 ms in width) by subframe division circuit 120. Speech signals divided into subframes are inputted into weighting filter 130. Weighting filter 130 performs perceptual weighting of inputted speech signals, and outputs the results to influence signal subtracter 140. Influence signal subtracter 140 subtracts the weighted influence signal from the preceding subframe, and outputs the result to candidate selection circuit 150 for the adaptive codebook, to candidate selection circuit 160 for the first excitation codebook, to candidate selection circuit 170 for the

second excitation codebook, and to gain codebook search circuit 210.

Subsequently, candidate selection circuit 150 for the adaptive codebook selects the candidate of L_0 pieces of adaptive codevectors from adaptive codebook 175 according to equation (3). Candidate selection circuit 150 for the adaptive codebook outputs the weighted synthetic signal of the candidate of the selected adaptive codevectors and delay d which constitutes the index of the candidate of adaptive codevectors, to candidate selection circuits 160, 170 for the first and second excitation codebooks and to optimum combination search circuit 200.

Candidate selection circuit 160 for the first excitation codebook selects the candidate of L_1 pieces of the first excitation codevector from first excitation codebook 180, according to the output of the influence signal subtracter, the spectral parameter and the candidate of the adaptive codevector by using equation (6). Candidate selection circuit 160 for the first excitation codebook outputs the weighted synthetic signal and index of the candidate of the selected first excitation codevector to optimum combination search circuit 200. In the same manner, candidate selection circuit 170 for the second excitation codebook selects the candidate of the second excitation codevector from the second excitation codebook according to equation (9), and outputs the weighted synthetic signal and index of the selected second excitation codevector to optimum combination search circuit 200.

Optimum combination search circuit 200 searches for the combination of the optimum candidates according to equation (14), and outputs the delay of the adaptive codevector and the indices of the first and second excitation codevectors to multiplexer 230, and weighted synthetic signals of each codevector to gain codebook search circuit 210. Gain codebook search circuit 210 searches for the optimum gain codevector from gain codebook 220 according to each of the inputted weighted synthetic signals, and outputs the index of thus obtained gain codevector to multiplexer 230.

Finally, multiplexer 230 assembles and outputs the code for the speech signal divided into subframes according to the delay and index outputted from optimum combination search circuit 200 and to the index outputted from gain codebook search circuit 210. By carrying out the above process, speech coding of every subframe is completed.

According to the present embodiment, the candidates are selected first from the adaptive codebook and each of excitation codebooks, and then the optimum combination is selected from the combination of each of thus selected candidates, so that a sufficiently good speech quality can be obtained with a relatively small operation. In addition, since the gain codebook which stores the quantized gain vectors is used for selecting the optimum combination from combinations of the candidates, SN ratio is further improved.

The second embodiment of the present invention will be described with reference to Fig. 2. In the voice-coder shown in Fig. 2, each block attached with the same reference numeral as that in Fig. 1 has the same function as that in Fig. 1.

The voice-coder in Fig. 2, when compared with the voice-coder in Fig. 1, differs in that it has no gain codebook search circuit and optimum combination search circuit, but has instead gain-including optimum combination search circuit 300. Gain-including optimum combination search circuit 300 receives candidates of the adaptive codevectors, candidates of the first and second excitation codevectors, and outputs of influence signal subtracter 140, and selects the optimum combination from all of the combinations of the candidates and gain codevectors by searching for gain codebook 220. Gain-including optimum combination search circuit 300 is structured so as to output the delay or index of each codevector composing the selected combination to multiplexer 230 according to the selected combination.

The search algorithm which controls gain-including optimum combination search circuit 300 will next be described.

Gain-including optimum combination search circuit 300 searches for the combination of candidates which has the minimum value of error E by calculating E for all of the combinations of candidates of the selected adaptive codevectors, the selected first and second excitation codevectors, and all of the gain codevectors, where E is calculated by the following equation:

$$E = \|z - Q\beta_k sa_d - Q\gamma_k se_i^1 - Q\delta_k se_j^2\|^2 \qquad (16)$$

where $Q\beta_k$, $Q\gamma_k$, $Q\delta_k$ are each gain codevector.

It is acceptable to use, in place of above $Q\beta_k$, $Q\gamma_k$, $Q\delta_k$, not the gain codevector itself, but gain codevectors converted by the matrix to be calculated from the quantized power of the weighted input signal, the weighted synthetic signal of the adaptive codevector and the weighted synthetic signals of the first and second excitation codevectors. Since it requires large operation to search for the minimum value of E by calculating it against all the gain codevectors, it is also possible to perform a preliminary selection of the gain codebook to reduce the operation. The preliminary selection of the gain codebook is performed, for

example, by selecting the predetermined fixed number of gain codevectors whose first components are close to the sequential optimum gain of the adaptive codevector.

The operation of this voice-coder will be described. It is the same as that of the voice-coder shown in Fig. 1 except that the candidates of vectors are outputted from each of candidate selection circuits 150, 160 and 170. These candidates of codevectors are inputted into gain-including optimum combination search circuit 300, whereby the optimum combination of candidates is searched according to equation (16). Then consulting the searched combination, the delay of the adaptive codevector and indices of the first and second excitation codevectors and gain codevectors are inputted into multiplexer 230, from which speech signal codes are outputted.

Next, the third embodiment of the present invention will be described with reference to Fig. 3. In the voice-coder shown in Fig. 3, each block attached with the same reference numeral as that in Fig. 1 has the same function as that in Fig. 1.

This voice-coder differs from the one shown in Fig. 1 in that the second excitation codebook is composed of excitation super codebook 390. A super codebook means a codebook which stores codevectors with the number of bits larger than the number of bits to be transmitted. Index i of the candidate of the first excitation codevector is outputted from first excitation codebook selection circuit 160 to second excitation super codebook 390. The selection of the candidate of the second excitation codevectors from second excitation super codebook 390 is carried out by searching codevectors from a portion of second excitation super codebook 390, the portion being expressed by set $F_2(i)$ of indices to be determined according to index i of the first excitation codevector.

When searching of the candidates of the first and second codevectors is finished, then the optimum combination of candidates is searched in optimum combination search circuit 200 according to equation (14) as searched in the first embodiment. In the present embodiment, it is possible to modify so as to output all the second excitation codevectors which correspond to set of indices $F_2(i)$ without performing selection of candidates of the second excitation codevectors in candidate selection circuit 170 of the second excitation codebook. In this case, optimum combination search circuit 200 can search the optimum combination from the combination of the candidate of the adaptive codevectors, the candidate of the first excitation codevectors, and all of the second excitation codevectors corresponding to set $F_2(i)$.

As described above in the third embodiment of the present invention, by applying the super codebook in the embodiment, it becomes possible to obtain speech quality as substantially good as the case with a excitation codebook of an increased codebook size without increasing the bit rates.

The fourth embodiment of the present invention will nest be described with reference to Fig. 4. In the voice-coder shown in Fig. 4, each block attached with the same reference numeral as that in Fig. 2 has the same function as that in Fig. 2.

This voice-coder uses second excitation super codebook 390 instead of the second excitation codebook, differently from the voice-coder in Fig. 2. Super codebook 390 is similar to the super codebook in the voice-coder shown in Fig. 3. The candidate of the second excitation codevector to be selected from second excitation super codebook 390 is also selected in the same way as in the third embodiment, and other operations are conducted in the same manner as in the second embodiment. In this case, it is also possible to modify candidate selection circuit 170 for the second excitation codevectors so as to output all of the second excitation codevectors which correspond to set of indices $F_2(i)$ without selecting the candidate of the second excitation codevectors.

Although each embodiment of the present invention has been described above, the operation of each embodiment can be modified in such a way that auto-correlation $\langle se_i, se_i \rangle$ of weighted synthetic signal se_i of the excitation codevector is obtained according to the following equation for the purpose of reducing the operation:

$$\langle se_i, se_i \rangle = hh(0)ee_i(0) + 2\sum_{i=1}^{im} hh(1)ee_i(1) \text{------}(17)$$

where hh is an auto-correlation function of the impulse response of a weighting synthesis filter, ee_i an auto-correlation function of the excitation code vector with index i, and im a length of the impulse response.

As well, cross-correlation between weighted synthetic signal se_i of the excitation codevector and arbitrary vector v can be calculated according to the following equation to reduce the operation:

$$\langle v, se_i \rangle = \langle H^T v, e_i \rangle \quad (18)$$

where H is an impulse response matrix of the weighting synthesis filter.

Cross-correlation between weighted synthetic signal sa_d of the adaptive codevector and arbitrary vector v can be obtained according to the following equation in the same way:

$$\langle v, sa_d \rangle = \langle H^T v, a_d \rangle \quad (19)$$

Further, in the case of searching for the optimum combination in the optimum combination search circuit of the first and third embodiments, although a particular limitation ($\gamma = \delta$) is now assigned to gains γ, δ of the first and second excitation codevectors as described above, it is possible to provide limitations other than $\gamma = \delta$ or to provide no limitation.

Further, it is also possible to apply a delayed decision system in each embodiment in such a way that the combination of candidates is selected so as to have the minimum cumulative error for the whole frames without uniquely determining the adaptive codevector, the first and second excitation codevectors and the gain codevector for each subframe while leaving the candidates undetermined.

It is to be understood that variations and modifications of the method for speech coding and of the voice-coder disclosed herein will be evident to those skilled in the art. It is intended that all such modifications and variations be included within the scope of the appended claims.

Claims

1. A method for speech coding for coding speech signals divided into frames spaced with a constant interval, wherein,

   an adaptive codebook storing excitation signals determined in the past and a plurality of excitation codebooks for multi-stage vector quantization of an excitation signal of the input speech signal are prepared;

   a spectral parameter of said input speech signal is obtained;

   said frame is divided into subframes;

   a candidate of a first fixed number of adaptive codevectors is selected for every said subframe from said adaptive codebook by using said input speech signal and said spectral parameter;

   candidates of a second fixed number of excitation codevectors are selected for every said subframe from said excitation codebooks, respectively, by using said input speech signal, said spectral parameter and the candidate of said adaptive codevector; and

   a combination of the adaptive codevector and each of the excitation codevectors forming an excitation signal of said subframe is selected from the candidates of said adaptive codevector and each of said excitation codevectors by using said input speech signal and said spectral parameter.

2. A method for speech coding according to Claim 1, wherein selection of the candidates of the adaptive codevector and each of the excitation codevectors are performed, respectively, in order, from the selection of the candidate with a smaller error.

3. A method for speech coding according to Claim 1 or 2, wherein,

   a gain codebook is used for performing quantization of gains of said adaptive codebook and each of said excitation codebooks, respectively; and

   a gain codevector is determined by using said gain codebook when selection of a combination of the adaptive codevector and each of the excitation codevectors forming the excitation signal of said subframe from the candidates of said adaptive codevector and said excitation codevector is performed.

4. A method for speech coding according to any of Claims 1 to 3, wherein,

   at least one or more of excitation super codebooks is included in said plurality of excitation codebooks, said super codebook comprising bits with the number of bits larger than the number of bits to be transmitted; and

   selection of the candidate of the excitation codevector from said excitation super codebook is performed corresponding to the candidate of the excitation codevector already selected.

5. A method for speech coding according to Claim 1, wherein the step of selecting the combination of the adaptive codevector and each of the excitation codevectors forming the excitation signal of said subframe from the candidates of said adaptive codevector and said excitation codevector, further

comprising the steps of:

determining the optimum gain codevector from said gain codebook; and

reflecting said gain codevector on said adaptive codevector and each of said excitation codevectors forming said excitation signal.

6. A method for speech coding according to Claim 5, wherein,

at least one or more of excitation super codebooks is included in said plurality of excitation codebooks, said super codebook comprising bits with the number of bits larger than the number of bits to be transmitted; and

selection of the candidate of the excitation codevector from said excitation super codebook is performed corresponding to the candidate of the excitation codevector already selected.

7. A voice-coder for coding speech signals divided into frames spaced with a constant interval, comprising:

linear prediction analysis means for outputting spectral parameters of input speech signals;

an adaptive codebook for storing excitation signals determined in the past;

a plurality of excitation codebooks provided for multi-stage vector quantization of the excitation signal of said input speech signals;

wherein, in case of searching for a combination of the adaptive codevector and each of the excitation codevectors for every subframe prepared by further division of said frame from said adaptive codebook and each of said excitation codebooks, respectively, said combination of the adaptive codevector and each of the excitation codevectors forming an excitation signal of said subframe:

a candidate of a first predetermined number of adaptive codevectors is selected from said adaptive codebook by using said input speech signal and said spectral parameter;

candidates of each predetermined number of excitation codevectors are selected from a plurality of said excitation codebooks, respectively, by using said input speech signal, said spectral parameter and the candidate of said adaptive codevector; and

a candidate of said adaptive codevector and each of said excitation codevectors forming the excitation signal of said subframe is selected from the candidate of said adaptive codevector and each of said excitation codevectors by using said input speech signal and said spectral parameter.

8. A voice-coder according to Claim 7, further comprising:

a gain codebook for quantization of each gain of said adaptive codebook and each of said excitation codebooks; wherein

said input speech signal, said spectral parameter and said gain codebook are used for searching a combination of the adaptive codevector and each of the excitation codevectors which forms the excitation signal of said subframe, from the candidates of said adaptive codevector and said excitation codevectors.

9. A voice-coder according to Claim 7 or 8, wherein,

at least one or more of excitation super codebooks is included in said plurality of excitation codebooks, said super codebook comprising bits with the number of bits larger than the number of bits to be transmitted; and

selection of the candidate of the excitation codevector from said excitation super codebook is performed corresponding to the candidate of the excitation codevector already selected.

10. A voice-coder for coding speech signals divided into frames spaced with a constant interval, comprising:

linear prediction analysis means for outputting spectral parameters of input speech signals;

an adaptive codebook storing excitation signals determined in the past;

a plurality of excitation codebooks provided for multi-stage vector quantization of an excitation signal of said input speech signals;

subframe division means for generating subframe signals by dividing said frame into subframes;

first selection means for selecting a candidate of a first fixed number of adaptive codevectors from said adaptive codebook in accordance with said subframe signal and said spectral parameter;

second selection means provided for every said excitation codebook for selecting the candidate of the excitation codevectors of the number predetermined for every excitation codebook, from the corresponding excitation codebook in accordance with said subframe signal, said spectral parameter

and the candidate of said adaptive codevector; and

means for searching the candidate of said adaptive codevector and each of said excitation codevectors which forms the excitation signal of said subframe, from the candidate of said adaptive codevector and the candidate of each of said excitation codevectors in accordance with said input speech signal and said spectral parameter.

11. A voice-coder according to Claim 10, wherein,

first and second selecting means select each corresponding candidate, in order, from the candidate with a smaller error;

said search means searches the candidate of said codevector on the condition of whose error is lowest.

12. A voice-coder according to Claim 10, further comprising:

a gain codebook for quantization of each gain of said adaptive codebook and each of said excitation codebooks; wherein

said search means searches the candidate of said codevector by further consulting said gain codebook.

13. A voice-coder according to Claim 11, further comprising:

a gain codebook for quantization of each gain of said adaptive codebook and each of said excitation codebooks; wherein

said search means further determines the optimum gain codevector from said gain codebook by consulting said gain codebook, and reflects said gain codevector on the adaptive codevector and each of the excitation codevector which forms said excitation signal.

14. A voice-coder according to any of Claims 10 to 13, wherein, at least one or more of excitation super codebooks is included in said plurality of excitation codebooks, said super codebook comprising bits with the number of bits larger than the number of bits to be transmitted; and

said second selection means corresponding to said excitation super codebook performs selection of the candidate of the excitation codevector from said excitation super codebook according to the candidate of the excitation codevector already selected.

FIG. 1

FIG. 2

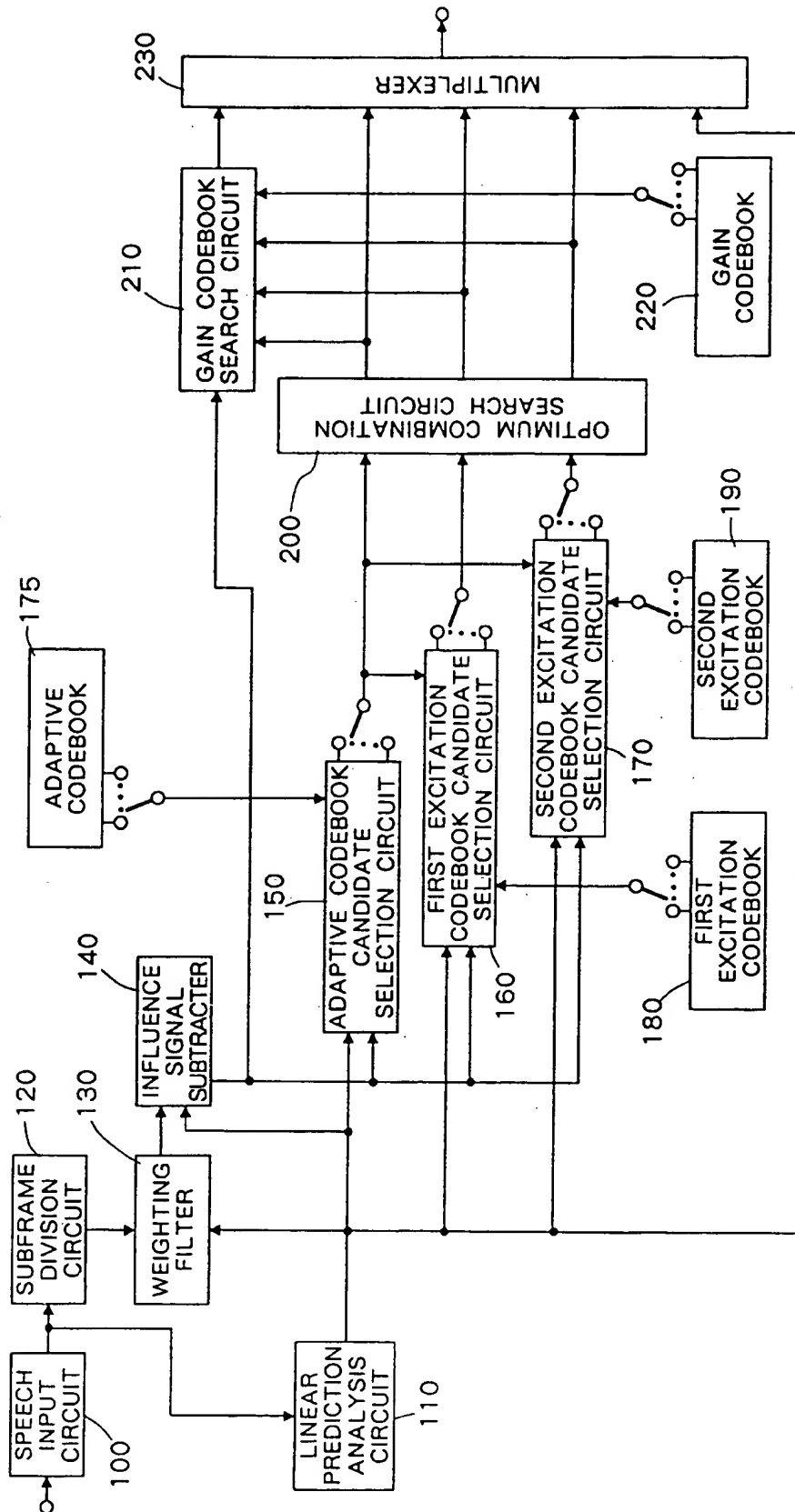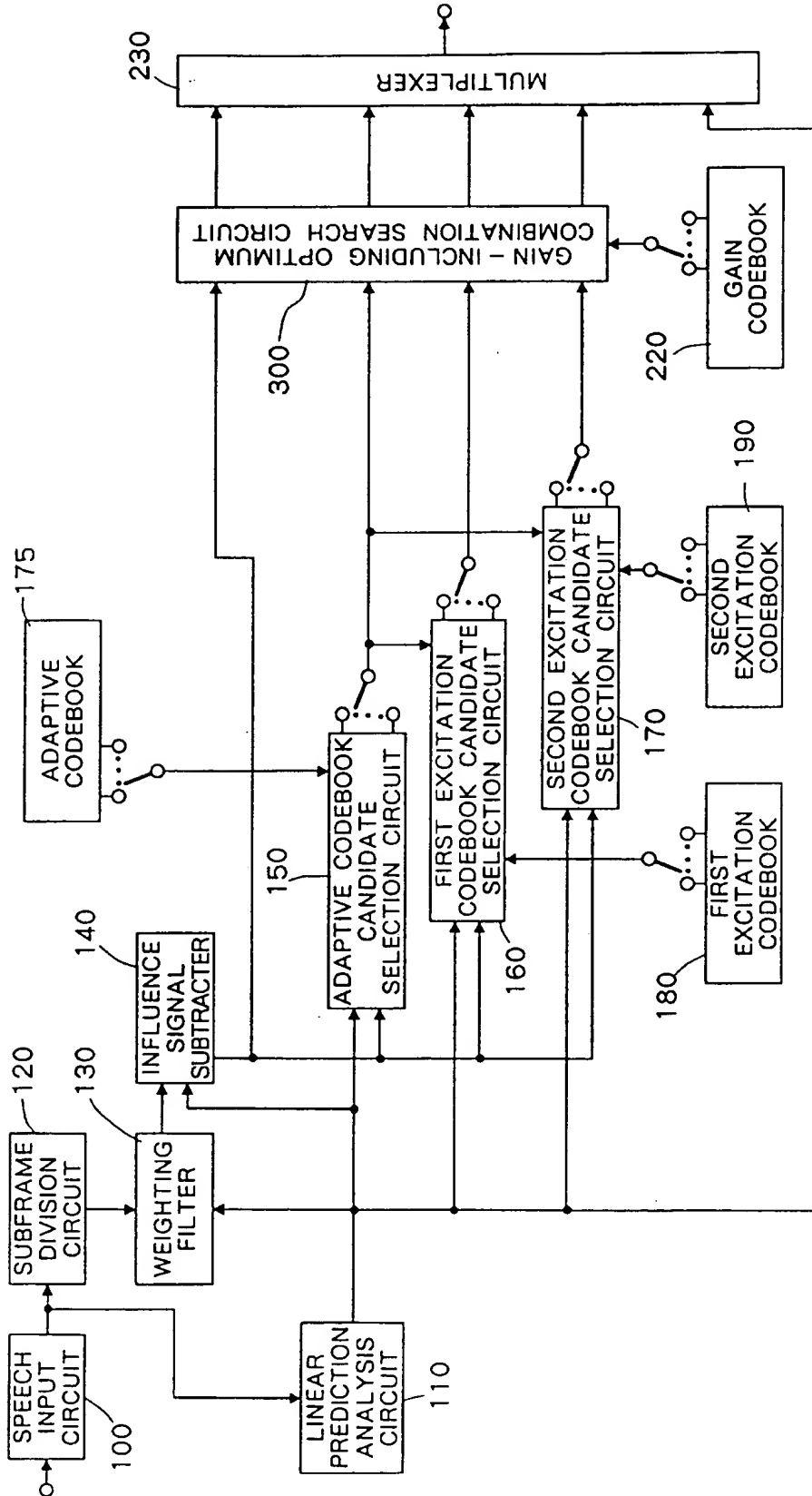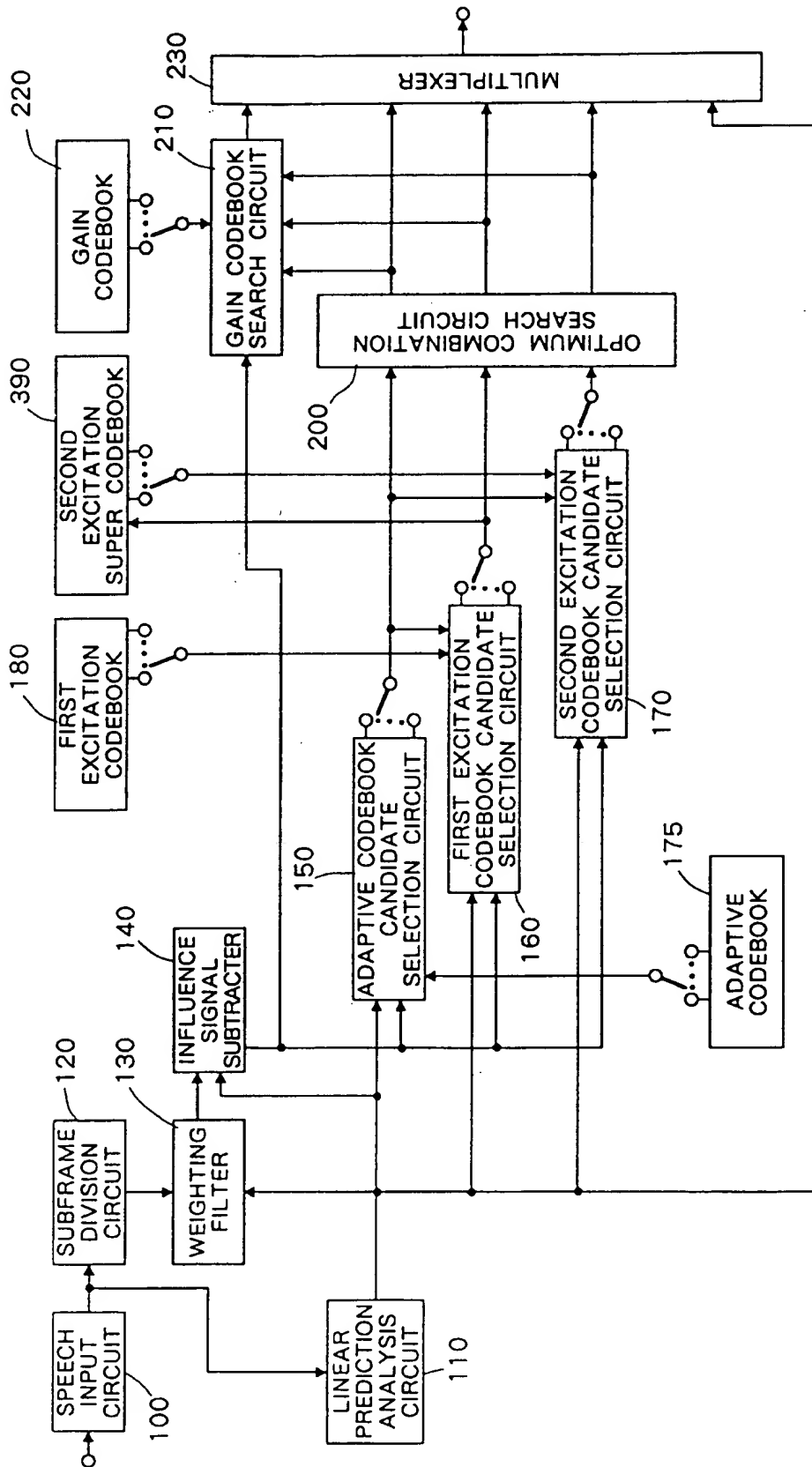FIG. 3

FIG. 4

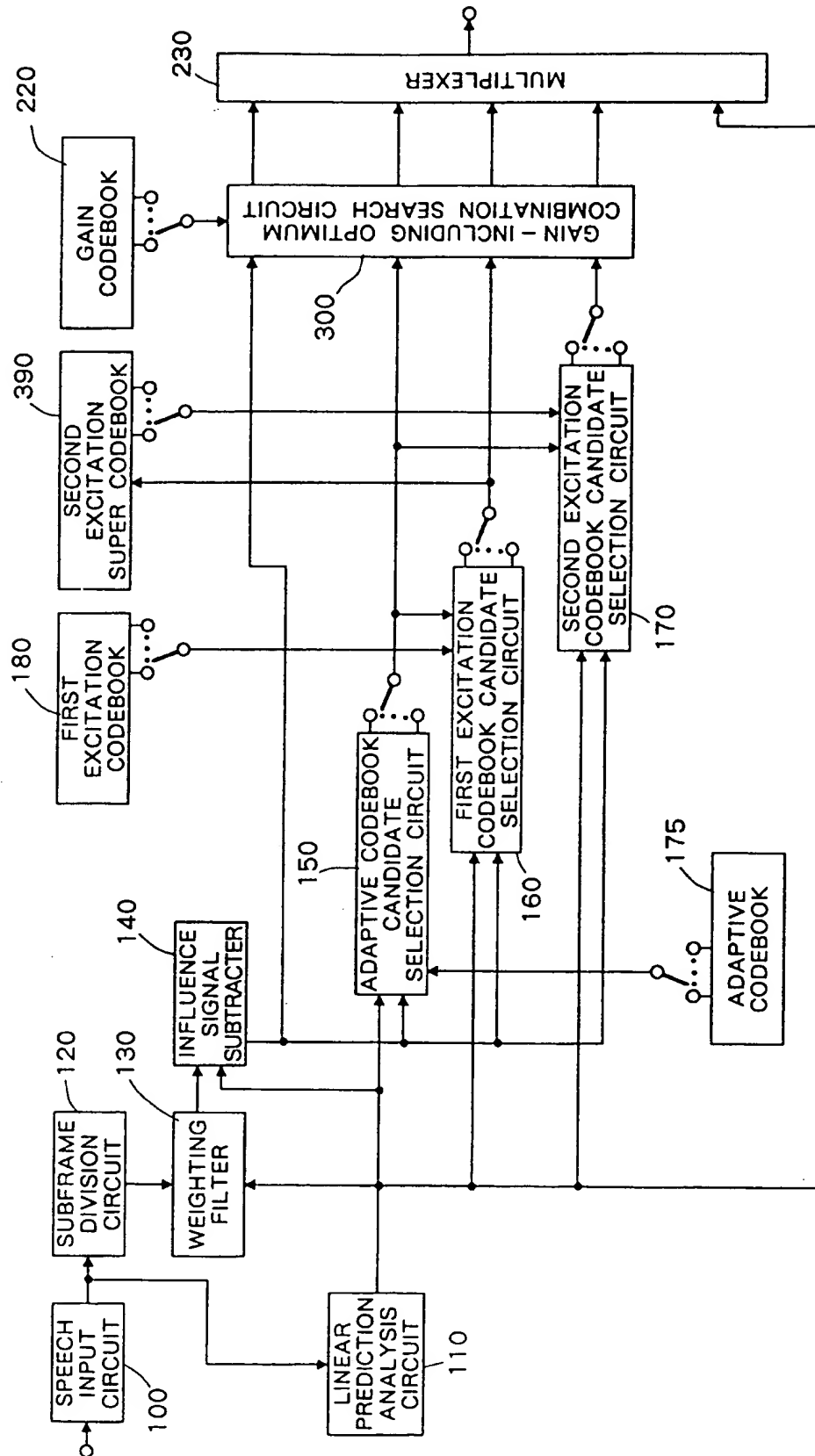## EUROPEAN PATENT APPLICATION

(21) Application number: 92120573.8

(51) Int. Cl.⁵: **G10L 9/14**

(22) Date of filing: 02.12.92

(54) **Method for speech coding and voice-coder.**

(57) A method for speech coding and a voice-coder
for coding speech signals divided into frames
spaced with a constant interval are disclosed. An
adaptive codebook storing excitation signal deter-
mined in advance and a plurality of excitation
codebooks for multi-stage vector quantization are
provided. Each frame is divided into subframes. For
each subframe, a candidate of a first predetermined
number of adaptive codevectors is selected, and
then candidates of each predetermined number of
excitation codevectors are selected from each ex-
citation codebook, respectively, by using the can-
didate of the adaptive codevector. Finally, a com-
bination of the adaptive codevector and each of the
excitation codevector is selected from the candi-
dates of the adaptive codevector and each of the
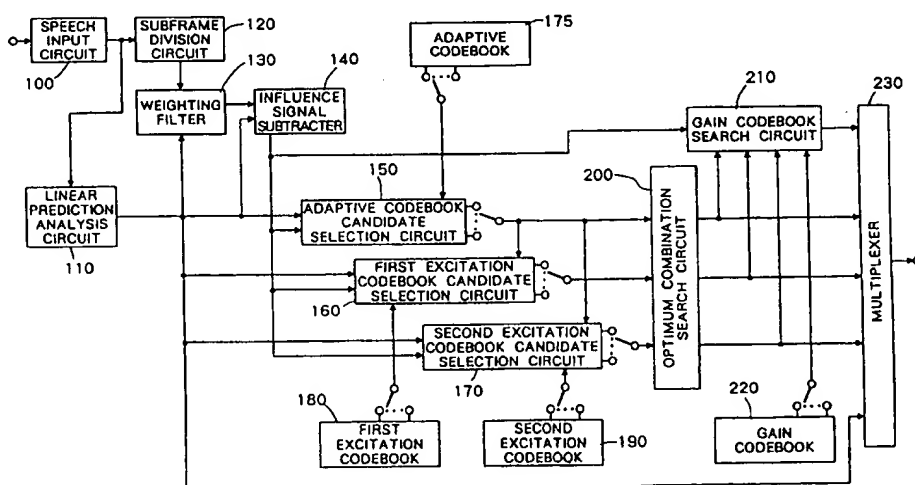sound codevectors.

FIG. 1

EP 0 545 386 A3

European Patent
Office

EUROPEAN SEARCH REPORT

Application Number

EP   92 12 0573
Page 1

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. Cl.5)
X	EP-A-0 443 548 (NEC)	1,2,7, 10,11	G10L9/14
A	* page 7, line 15 - page 10, line 16; figure 1 * * page 13, line 16 - page 14, line 23; claims 1-4,6; figure 5 *	3	
P,X	EP-A-0 477 960 (NEC) * page 2, line 54 - page 4, line 52; figure 1 *	1,2,7, 10,11	
A	NEC RESEARCH AND DEVELOPMENT vol. 32, no. 4, October 1991, TOKYO JAPAN pages 549 - 555 UNNO ET AL 'Learned codebook excited linear predictive speech codec for digital cellular system' * abstract * * page 550-551, section 2,3; figure 1; table II *	1,2,4	
A	IEEE TRANSACTIONS ON CONSUMER ELECTRONICS vol. 37, no. 4, November 1991, NEW YORK US pages 772 - 782 SUNWOO , PARK 'Real time implementation of the VSELP on a 16-bit DSP chip' * pages 772-773, section IIA, figures 1, 2 *	1	TECHNICAL FIELDS SEARCHED (Int. Cl.5) G10L
A	INTERNATIONAL CONFERENCE ON ACOUSTICS SPEECH AND SIGNAL PROCESSING vol. 1, 23 May 1989, GLASGOW SCOTLAND U.K. pages 65 - 68 JAYANT, CHEN 'Speech coding with time varying bit allocations to excitation and LPC parameters' * Introduction; table 3; figure 2 *	4,9	

-/--

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	22 JUNE 1993	FARASSOPOULOS A.

EPO FORM 1503 03.82 (P0401)

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. Cl.5)
A	INTERNATIONAL CONFERENCE ON ACOUSTICS SPEECH AND SIGNAL PROCESSING vol. 1, 3 April 1990, ALBUQUERQUE NEW MEXICO USA pages 461 - 464 GERSON, JASIUK 'Vector sum excited linear prediction speech coding at 8Kbps' * pages 461-463, sections I-V *	1-3,7,8, 10-12	

			TECHNICAL FIELDS SEARCHED (Int. Cl.5)

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	22 JUNE 1993	FARASSOPOULOS A.

EPO FORM 1503 03.82 (P0401)

(12)   EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention
of the grant of the patent:
07.04.1999 Bulletin 1999/14

(51) Int Cl.^{6}: **G10L 9/14**

(21) Application number: 92120573.8

(22) Date of filing: **02.12.1992**

(54) **Method for speech coding and voice-coder**

Methode zur Sprachkodierung und Sprachkodierer

Méthode pour le codage de la parole et codeur de parole

(84) Designated Contracting States:
DE FR GB

(30) Priority: **03.12.1991 JP 319314/91**

(43) Date of publication of application:
09.06.1993 Bulletin 1993/23

(73) Proprietor: **NEC CORPORATION**
Tokyo (JP)

(72) Inventor: **Miyano, Toshiki**
Minato-ku, Tokyo (JP)

(74) Representative: **VOSSIUS & PARTNER**
Postfach 86 07 67
81634 München (DE)

(56) References cited:
EP-A- 0 443 548   **EP-A- 0 477 960**

• NEC RESEARCH AND DEVELOPMENT vol. 32,
no. 4, October 1991, TOKYO JAPAN pages 549 -
555 UNNO ET AL 'Learned codebook excited
linear predictive speech codec for digital cellular
system'
• IEEE TRANSACTIONS ON CONSUMER
ELECTRONICS vol. 37, no. 4, November 1991,
NEW YORK US pages 772 - 782 SUNWOO, PARK
'Real time Implementation of the VSELP on a
16-bit DSP chip'
• INTERNATIONAL CONFERENCE ON
ACOUSTICS SPEECH AND SIGNAL
PROCESSING vol. 1, 23 May 1989, GLASGOW
SCOTLAND U.K. pages 65 - 68 JAYANT, CHEN
'Speech coding with time varying bit allocations
to excitation and LPC parameters'
• INTERNATIONAL CONFERENCE ON
ACOUSTICS SPEECH AND SIGNAL
PROCESSING vol. 1, 3 April 1990,
ALBUQUERQUE NEW MEXICO USA pages 461 -
464 GERSON, JASIUK 'Vector sum excited linear
prediction speech coding at 8Kbps'

EP 0 545 386 B1

Description

[0001] The present invention relates to a method for speech coding and to a voice-coder, particularly to a method for speech coding and to a voice-coder which can achieve high coding quality with a relatively small operation at bit rates not greater than 8 kbit/s.

[0002] As a speech coding system to be applied to vector quantization of excitation signals at low bit rates by using a excitation codebook comprising random numbers, a CELP system described in a paper (hereinafter referred to as literature 1) titled "CODE-EXCITED LINEAR PREDICTION (CELP): HIGH-QUALITY SPEECH AT VERY LOW BIT RATES" (Proc. ICASSP, pp. 937-940, 1985) by Manfred R. Shroeder and Bishnu S. Atal is known. There is also a CELP system using an adaptive codebook described in a paper (hereinafter referred to as literature 2) titled "IM-PROVED SPEECH QUALITY AND EFFICIENT VECTOR QUANTIZATION IN CELP" (Proc. ICASSP, pp. 155-158, 1988) by W. B. Kleijin, D. J. Krasinski and R. H. Ketchum. The CELP system using the adaptive codebook receives speech signals divided into frames spaced with a constant interval. The CELP utilizes a linear predictive analyzer for obtaining spectral parameters of input speech signals, the adaptive codebook having excitation signals determined in the past, the excitation codebook comprising random numbers to be used for vector quantization of the excitation signals of said input speech signals. The CELP selects an adaptive codevector by using the input speech signal and the synthesized signal of the adaptive codevector for every subframe made by equally dividing the frame. Subsequently, CELP performs selection of excitation codevectors by using the input signals, the synthesized signal of the selected adaptive codevector and said excitation codevector.

[0003] However, the CELP systems have the following disadvantage, in that a quite large operation is required for searching the excitation codebook. Moreover since the adaptive codebook is determined independently of the excitation codebook, it is impossible to get a high SN (signal to noise) ratio. Further in the above CELP system, although the adaptive codebook and the excitation codebook are each searched by using gains not quantized, it becomes impossible to obtain a higher SN ratio when the adaptive codebook and the excitation codebook are searched for all the quantization value of gains. Furthermore, it is impossible to obtain sufficiently good speech quality with low bit rates such as 8 kbit/s or less because of the too small size of the excitation codebook.

[0004] A speech coder has been described in EP-A-0 443 548. The speech coder particularly comprises an LPC analyzer, a difference signal generating section, a first code book, a second code book, and a multiplexer. The LPC analyzer divides an input discrete speech signal into signal components in units of frames each having a predetermined time length, and obtains a spectrum parameter representing a spectrum envelope of the speech signal. The difference signal generating section obtains a difference signal by dividing the frame into subframes each having a predetermined time length, and predicting a pitch parameter representing a long-term correlation on the basis of a past sound source signal.

The multiplexer represents a sound source signal of the speech signal by a linear combination of a signal selected from the first code book in accordance with each obtained difference signal and a signal selected from the second code book, and outputs the combination.

[0005] A coder and a decoder for speech signals have been described in NEC Research & Development, vol. 32, no. 4, October 1991, Tokyo, Japan, pages 549 - 555, Unno et al., "Learned Codebook Excited Linear Predictive (LCELP) Speech Codec for Digital Cellular System".

[0006] An object of the present invention is to provide a method for speech coding which can solve the above problem of the conventional method and achieve high quality speech by a relatively small operation even at low bit rates such as less than 8 kbit/s.

[0007] Another object of the present invention is to provide a voice-coder which can solve the above problem of the conventional method and achieve high quality speech by a relatively small operation even at low bit rates such as less than 8 kbit/s.

[0008] These objects are achieved by the features of claims 1 and 7, respectively.

[0009] The above and other objects, features and advantages of the present invention will be apparent from the following description referring to the accompanying drawings which illustrate examples of preferred embodiments of the present invention.

BRIEF DESCRIPTION OF THE DRAWINGS:

[0010] Fig. 1 is a block diagram showing the structure of a voice-coder of a first embodiment of the present invention.

[0011] Fig. 2 is a block diagram showing the structure of a voice-coder of a second embodiment of the present invention.

[0012] Fig. 3 is a block diagram showing the structure of a voice-coder of a third embodiment of the present invention.

[0013] Fig. 4 is a block diagram showing the structure of a voice-coder of a fourth embodiment of the present invention.

DESCRIPTION OF THE PREFERRED EMBODIMENTS:

[0014] A first preferable embodiment of the present invention will be described with reference to Fig. 1. In the voice-coder shown in Fig.1, there are provided adaptive codebook 175, gain codebook 220 and two kinds of excitation codebooks 180, 190.

[0015] Speech input circuit 100 is provided for receiving speech signals divided into frames spaced with a constant interval. Subframe division circuit 120 and linear prediction analysis circuit 110 are provided on the output side of speech input circuit 100. Subframe division circuit 120 outputs subframes by equally dividing the frame, and linear prediction analysis circuit 110 performs linear prediction analyses of speech signals composing frames for obtaining spectral parameters of the speech signals. On the output side of subframe division circuit 120, weighting filter 130 is provided for performing perceptual weighting by receiving subframes and spectral parameters. On the output side of weighting filter 130, influence signal subtracter 140 is provided for subtracting weighted influence signal from the preceding subframe and outputting the results thereof.

[0016] Adaptive codebook 175 stores excitation signals decided in the past as adaptive codevectors. Corresponding to adaptive codebook 175, adaptive codebook candidate selection circuit 150 is provided for selecting the previously fixed number of adaptive codevectors and for outputting thereof as candidates of adaptive codevectors. Adaptive codebook candidate selection circuit 150 performs selection of the candidate according to the spectral parameter and the output signal of influence signal subtracter 140.

[0017] First and second excitation codebooks 180, 190 operate for multi-stage vector quantization of the excitation signal, and store the first and second excitation codevectors, respectively. Corresponding to first and second excitation codebooks 180, 190, candidate selection circuits 160, 170 for the first and second excitation codebooks are provided respectively. Candidate selection circuits 160, 170 select the previously fixed number of excitation codevectors from corresponding respective excitation codebooks 180, 190 and output them as the candidates of the excitation codevectors. Spectral parameters, output signals of the influence signal subtracter and candidates of adaptive codevectors are inputted into each of candidate selection circuits 160, 170 for the excitation codebook.

[0018] Optimum combination search circuit 200 is provided for candidates selected by candidate selection circuits 150, 160, 170 for the corresponding codebooks in order to search the optimum combination of candidates. Further, gain codebook search circuit 210 and multiplexer 230 are provided. Optimum combination search circuit 200 is structured so as to output to multiplexer 230 the delay (to the adaptive codevector) or index (to each excitation codevector) with reference to each of the respective optimum adaptive codevectors, to the first and second excitation codevectors according to the results of the search, and to the output weighted synthetic signals of the above vectors to gain codebook search circuit 210, respectively. Gain codebook search circuit 210 searches for the optimum gain codevector from gain codebook 220 which stores gain codevectors, and outputs the index of thus searched optimum gain codevector. Multiplexer 230 is structured so as to receive delay or indices from optimum combination search circuit 200 or gain codebook search circuit 210, and output codes which correspond to input speech signals according to delay or indices.

[0019] Next, description will be made with reference to selection or search algorithm of each candidate selection circuit 150, 160, 170 or optimum combination search circuit 200 of the present embodiment. Under these algorithms, the excitation signal is processed by two-stage vector quantization by using two kinds of excitation codebooks 180, 190.

[0020] First, in adaptive codebook candidate selection circuit 150, the predetermined number L_0 of the adaptive codevectors is selected, in order, from the one with smaller error E_0 expressed by equation (1):

$$E_0 = \left\| z - \beta_0 sa_d \right\|^2 \tag{1}$$

where, z is a signal obtained by subtracting an influence signal from a perceptually weighted input signal, sa_d a perceptually weighted synthetic signal of adaptive codevector a_d with delay d, β_0 a sequential optimum gain of an adaptive codevector, $\| \ \|$ Euclid norm. The sequential optimum gain β_0 of the adaptive codevector is given by:

$$\beta_0 = \frac{<z, sa_d>}{<sa_d, sa_d>} \tag{2}$$

[0021] By substituting above equation (2) into equation (1), the next equation is obtained.

$$E_0 = \|z\|^2 - \frac{<z, sa_d>^2}{<sa_d, sa_d>} \tag{3}$$

where, $<,>$ represents an inner product.

[0022] In candidate selection circuit 160 for the first excitation codebook, candidates of predetermined number L_1 of the first excitation codevectors are selected for each L_0 piece of the adaptive codevectors selected by candidate selection circuit 150 for the adaptive codebook, in order, from the one with smaller error E_1 expressed by equation (4):

$$E_1 = \left\| za - \gamma_0 se_i^1 \right\|^2 \tag{4}$$

where se_i^1 is a perceptually weighted synthesized signal of first excitation codevector e_i^1 with index i, γ_0 a sequential optimum gain of the first excitation codevector, and $za = z - \beta_0 sa_d$.

[0023] Therefore:

$$\gamma_0 = \frac{<za, se_i^1>}{<se_i^1, se_i^1>} \tag{5}$$

[0024] By substituting above equation (5) into equation (4), equation (6) below is obtained:

$$E_1 = \|z\|^2 - \frac{<z, se_i^1>^2}{<se_i^1, se_i^1>} \tag{6}$$

[0025] In the same way as described above, in candidate selection circuit 170 for the second excitation codebook, the candidates of predetermined number L_2 of the second excitation codevectors are selected for each L_0 piece of adaptive codevectors selected by candidate selection circuit 150 for the adaptive codebook, in order, from the one with smaller error E_2 expressed by the next equation:

$$E_2 = \left\| za - \delta_0 se_j^2 \right\|^2 \tag{7}$$

where se_j^2 is a perceptually weighted synthesized signal of second excitation codevector e_j^2 with index j, and δ_0 a sequential optimum gain of the second excitation codevector. Therefore:

$$\delta_0 = \frac{<za, se_j^2>}{<se_j^2, se_j^2>} \tag{8}$$

[0026] By substituting equation (8) into equation (7), following equation (9) is obtained.

$$E_2 = \|za\|^2 - \frac{<z, se_j^2>^2}{<se_j^2, se_j^2>} \tag{9}$$

[0027] In optimum combination search circuit 200, error E is calculated by the following equation for all the combinations of candidates of the selected adaptive codevectors, and the first and second excitation codevectors, and then the combination of the candidates with minimum E is searched.

$$E = \left\| z - \beta sa_d - \gamma se_1^1 - \delta se_j^2 \right\|^2 \tag{10}$$

where β, γ, δ are simultaneous optimum gains of an adaptive codevector, the first and second excitation codevectors, respectively. Therefore:

4

$$\begin{pmatrix} \beta \\ \gamma \\ \delta \end{pmatrix} = R^{-1} \begin{pmatrix} <z,\ sa_d> \\ <z,\ se_i{}^1> \\ <z,\ sa_j{}^2> \end{pmatrix} \qquad (11)$$

[0028]   However, R is to satisfy the following equation:

$$R = \begin{pmatrix} <sa_d,sa_d> & <sa_d,se_i{}^1> & <sa_d,\ se_j{}^2> \\ <se_i{}^1,sa_d> & <se_i{}^1,se_i{}^1> & <se_i{}^1,se_j{}^2> \\ <se_j{}^2,sa_d> & <se_j{}^2,se_i{}^1> & <se_j{}^2,se_j{}^2> \end{pmatrix} \qquad ---------- (12)$$

[0029]   By substituting equation (11) into equation (10), thus obtained,

$$R = \|z\|^2 - (<z,sa_d>\ <z,se_i{}^1>\ <z,se_j{}^2>) \times$$
$$R^{-1} \begin{pmatrix} <z,\ sa_d> \\ <z,\ se_i{}^1> \\ <z,\ se_j{}^2> \end{pmatrix} \qquad ----------(13)$$

[0030]   When above error E is calculated, it is acceptable to assign a particular limitation to simultaneous optimum gains γ, δ of each excitation codevector. For example, with the limitation that γ and δ are equal, error E is given by,

$$E = \|z\|^2 - (<z,sa_d>\ <z,se_i{}^1 + se_j{}^2>) \times$$
$$R^{-1} \begin{pmatrix} <z,\ sa_d> \\ <z,\ se_i{}^1+se_j{}^2> \end{pmatrix} \qquad ----------(14)$$

where,

$$R = \begin{pmatrix} <sa_d,\ sa_d> & <sa_d,\ se_i{}^1 + se_j{}^2> \\ <se_i{}^1+se_j{}^2,\ sa_d> & <se_i{}^1+se_j{}^2,\ se_i{}^1+se_j{}^2> \end{pmatrix} \qquad ----------(15)$$

[0031]   Next, description will be made with reference to operation of the voice-coder of the present embodiment.

[0032]   Speech input circuit 100 receives speech signals divided into each frame (e.g., 40 ms in width), which signals are outputted to linear prediction analysis circuit 110 and subframe division circuit 120. In linear prediction analysis circuit 110, linear prediction analysis of the inputted speech signal is performed for calculating the spectral parameter. This spectral parameter is outputted to weighting filter 130, to influence signal subtracter 140, to candidate selection

circuit 150 for the adaptive codebook, to candidate selection circuit 160 for the first excitation codebook, to candidate selection circuit 170 for the second excitation codebook, and to multiplexer 230. Separately, a frame is divided into subframes (e.g., 8 ms in width) by subframe division circuit 120. Speech signals divided into subframes are inputted into weighting filter 130. Weighting filter 130 performs perceptual weighting of inputted speech signals, and outputs the results to influence signal subtracter 140. Influence signal subtracter 140 subtracts the weighted influence signal from the preceding subframe, and outputs the result to candidate selection circuit 150 for the adaptive codebook, to candidate selection circuit 160 for the first excitation codebook, to candidate selection circuit 170 for the second excitation codebook, and to gain codebook search circuit 210.

[0033]   Subsequently, candidate selection circuit 150 for the adaptive codebook selects the candidate of L_0 pieces of adaptive codevectors from adaptive codebook 175 according to equation (3). Candidate selection circuit 150 for the adaptive codebook outputs the weighted synthetic signal of the candidate of the selected adaptive codevectors and delay d which constitutes the index of the candidate of adaptive codevectors, to candidate selection circuits 160, 170 for the first and second excitation codebooks and to optimum combination search circuit 200.

[0034]   Candidate selection circuit 160 for the first excitation codebook selects the candidate of L_1 pieces of the first excitation codevector from first excitation codebook 180, according to the output of the influence signal subtracter, the spectral parameter and the candidate of the adaptive codevector by using equation (6). Candidate selection circuit 160 for the first excitation codebook outputs the weighted synthetic signal and index of the candidate of the selected first excitation codevector to optimum combination search circuit 200. In the same manner, candidate selection circuit 170 for the second excitation codebook selects the candidate of the second excitation codevector from the second excitation codebook according to equation (9), and outputs the weighted synthetic signal and index of the selected second excitation codevector to optimum combination search circuit 200.

[0035]   Optimum combination search circuit 200 searches for the combination of the optimum candidates according to equation (14), and outputs the delay of the adaptive codevector and the indices of the first and second excitation codevectors to multiplexer 230, and weighted synthetic signals of each codevector to gain codebook search circuit 210. Gain codebook search circuit 210 searches for the optimum gain codevector from gain codebook 220 according to each of the inputted weighted synthetic signals, and outputs the index of thus obtained gain codevector to multiplexer 230.

[0036]   Finally, multiplexer 230 assembles and outputs the code for the speech signal divided into subframes according to the delay and index outputted from optimum combination search circuit 200 and to the index outputted from gain codebook search circuit 210. By carrying out the above process, speech coding of every subframe is completed.

[0037]   According to the present embodiment, the candidates are selected first from the adaptive codebook and each of excitation codebooks, and then the optimum combination is selected from the combination of each of thus selected candidates, so that a sufficiently good speech quality can be obtained with a relatively small operation. In addition, since the gain codebook which stores the quantized gain codevectors is used for selecting the optimum combination from combinations of the candidates, SN ratio is further improved.

[0038]   The second embodiment of the present invention will be described with reference to Fig. 2. In the voice-coder shown in Fig. 2, each block attached with the same reference numeral as that in Fig. 1 has the same function as that in Fig. 1.

[0039]   The voice-coder in Fig. 2, when compared with the voice-coder in Fig. 1, differs in that it has no gain codebook search circuit and optimum combination search circuit, but has instead gain-including optimum combination search circuit 300. Gain-including optimum combination search circuit 300 receives candidates of the adaptive codevectors, candidates of the first and second excitation codevectors, and outputs of influence signal subtracter 140, and selects the optimum combination from all of the combinations of the candidates and gain codevectors by searching for gain codebook 220. Gain-including optimum combination search circuit 300 is structured so as to output the delay or index of each codevector composing the selected combination to multiplexer 230 according to the selected combination.

[0040]   The search algorithm which controls gain-including optimum combination search circuit 300 will next be described.

[0041]   Gain-including optimum combination search circuit 300 searches for the combination of candidates which has the minimum value of error E by calculating E for all of the combinations of candidates of the selected adaptive codevectors, the selected first and second excitation codevectors, and all of the gain codevectors, where E is calculated by the following equation:

$$E = \left\| z - Q\beta_k sa_d - Q\gamma_k se_i^1 - Q\delta_k se_i^2 \right\|^2 \qquad (16)$$

where $Q\beta_k$, $Q\gamma_k$, $Q\delta_k$ are each gain codevector.

[0042]   It is acceptable to use, in place of above $Q\beta_k$, $Q\gamma_k$, $Q\delta_k$, not the gain codevector itself, but gain codevectors

converted by the matrix to be calculated from the quantized power of the weighted input signal, the weighted synthetic signal of the adaptive codevector and the weighted synthetic signals of the first and second excitation codevectors. Since it requires large operation to search for the minimum value of E by calculating it against all the gain codevectors, it is also possible to perform a preliminary selection of the gain codebook to reduce the operation. The preliminary selection of the gain codebook is performed, for example, by selecting the predetermined fixed number of gain codevectors whose first components are close to the sequential optimum gain of the adaptive codevector.

[0043]   The operation of this voice-coder will be described. It is the same as that of the voice-coder shown in Fig. 1 except that the candidates of vectors are outputted from each of candidate selection circuits 150, 160 and 170. These candidates of codevectors are inputted into gain-including optimum combination search circuit 300, whereby the optimum combination of candidates is searched according to equation (16). Then consulting the searched combination, the delay of the adaptive codevector and indices of the first and second excitation codevectors and gain codevectors are inputted into multiplexer 230, from which speech signal codes are outputted.

[0044]   Next, the third embodiment of the present invention will be described with reference to Fig. 3. In the voice-coder shown in Fig. 3, each block attached with the same reference numeral as that in Fig. 1 has the same function as that in Fig. 1.

[0045]   This voice-coder differs from the one shown in Fig. 1 in that the second excitation codebook is composed of excitation super codebook 390. A super codebook means a codebook which stores codevectors with a number of bits larger than the number of bits to be transmitted. Index i of the candidate of the first excitation codevector is outputted from first excitation codebook selection circuit 160 to second excitation super codebook 390. The selection of the candidate of the second excitation codevectors from second excitation super codebook 390 is carried out by searching codevectors from a portion of second excitation super codebook 390, the portion being expressed by set $F_2(i)$ of indices to be determined according to index i of the first excitation codevector.

[0046]   When searching of the candidates of the first and second codevectors is finished, then the optimum combination of candidates is searched in optimum combination search circuit 200 according to equation (14) as searched in the first embodiment. In the present embodiment, it is possible to modify so as to output all the second excitation codevectors which correspond to set of indices $F_2(i)$ without performing selection of candidates of the second excitation codevectors in candidate selection circuit 170 of the second excitation codebook. In this case, optimum combination search circuit 200 can search the optimum combination from the combination of the candidate of the adaptive codevectors, the candidate of the first excitation codevectors, and all of the second excitation codevectors corresponding to set $F_2(i)$.

[0047]   As described above in the third embodiment of the present invention, by applying the super codebook in the embodiment, it becomes possible to obtain speech quality as substantially good as the case with a excitation codebook of an increased codebook size without increasing the bit rates.

[0048]   The fourth embodiment of the present invention will next be described with reference to Fig. 4. In the voice-coder shown in Fig. 4, each block attached with the same reference numeral as that in Fig. 2 has the same function as that in Fig. 2.

[0049]   This voice-coder uses second excitation super codebook 390 instead of the second excitation codebook, differently from the voice-coder in Fig. 2. Super codebook 390 is similar to the super codebook in the voice-coder shown in Fig. 3. The candidate of the second excitation codevector to be selected from second excitation super codebook 390 is also selected in the same way as in the third embodiment, and other operations are conducted in the same manner as in the second embodiment. In this case, it is also possible to modify candidate selection circuit 170 for the second excitation codevectors so as to output all of the second excitation codevectors which correspond to set of indices $F_2(i)$ without selecting the candidate of the second excitation codevectors.

[0050]   Although each embodiment of the present invention has been described above, the operation of each embodiment can be modified in such a way that auto-correlation $<se_i, se_i>$ of weighted synthetic signal se_i of the excitation codevector is obtained according to the following equation for the purpose of reducing the operation:

$$<se_i, se_i> = hh(0)ee_i(0) + 2\sum_{i=1}^{im} hh(1)ee_i(1) ------(17)$$

where hh is an auto-correlation function of the impulse response of a weighting synthesis filter, ee_i an auto-correlation function of the excitation code vector with index i, and im a length of the impulse response.

[0051]   As well, cross-correlation between weighted synthetic signal se_i of the excitation codevector and arbitrary vector v can be calculated according to the following equation to reduce the operation:

$$<v, se_i> = <H^T v, e_i>\qquad(18)$$

where H is an impulse response matrix of the weighting synthesis filter.

[0052] Cross-correlation between weighted synthetic signal sa_d of the adaptive codevector and arbitrary vector v can be obtained according to the following equation in the same way:

$$<v, sa_d> = <H^T v, a_d>\qquad(19)$$

[0053] Further, in the case of searching for the optimum combination in the optimum combination search circuit of the first and third embodiments, although a particular limitation ($\gamma = \delta$) is now assigned to gains γ, δ of the first and second excitation codevectors as described above, it is possible to provide limitations other than $\gamma = \delta$ or to provide no limitation.

[0054] Further, it is also possible to apply a delayed decision system in each embodiment in such a way that the combination of candidates is selected so as to have the minimum cumulative error for the whole frames without uniquely determining the adaptive codevector, the first and second excitation codevectors and the gain codevector for each subframe while leaving the candidates undetermined.

[0055] It is to be understood that variations and modifications of the method for speech coding and of the voice-coder disclosed herein will be evident to those skilled in the art. It is intended that all such modifications and variations be included within the scope of the appended claims.

Claims

1. A method for speech coding for coding speech signals divided into frames spaced with a constant interval, wherein:

   an adaptive codebook storing excitation signals determined in the past as adaptive codevectors and a plurality of excitation codebooks for multi-stage vector quantization of an excitation signal of the input speech signal are prepared;
   a spectral parameter of said input speech signal is obtained;
   said frame is divided into subframes; characterised in that
   a first fixed number (L0) of candidates of adaptive codevectors is selected for every said subframe from said adaptive codebook by using said input speech signal and said spectral parameter;
   a fixed number (L1, L2) of candidates of excitation codevectors are selected for every said subframe from each of said excitation codebooks, respectively, by using said input speech signal, said spectral parameter and said candidates of said adaptive codevectors; and
   an optimum combination of the adaptive codevector and each of the excitation codevectors forming the excitation signal of said subframe is selected from all the combinations of said candidates of said adaptive codevectors and said candidates of each of said excitation codevectors by using said input speech signal and said spectral parameter, the optimum combination being the one which minimizes a predetermined error function.

2. A method for speech coding according to Claim 1, wherein selection of the candidates of the adaptive codevector and each of the excitation codevectors are performed, respectively, in order, from the selection of the candidate with a smaller error.

3. A method for speech coding according to Claim 1 or 2, wherein,

   a gain codebook is used for performing quantization of gains of said adaptive codebook and each of said excitation codebooks, respectively; and
   a gain codevector is determined by using said gain codebook when selection of a combination of the adaptive codevector and each of the excitation codevectors forming the excitation signal of said subframe from the candidates of said adaptive codevector and said excitation codevector is performed.

4. A method for speech coding according to any of Claims 1 to 3, wherein:

   at least one or more of excitation super codebooks is included in said plurality of excitation codebooks, said

super codebook comprising bits with a number of bits larger than the number of bits to be transmitted; and selection of the candidate of the excitation codevector from said excitation super codebook is performed corresponding to the candidate of the excitation codevector already selected.

5. A method for speech coding according to Claim 1, wherein the step of selecting the combination of the adaptive codevector and each of the excitation codevectors forming the excitation signal of said subframe from the candidates of said adaptive codevector and said excitation codevector, further comprising the steps of:

   determining the optimum gain codevector from said gain codebook; and
   reflecting said gain codevector on said adaptive codevector and each of said excitation codevectors forming said excitation signal.

6. A method for speech coding according to Claim 5, wherein:

   at least one or more of excitation super codebooks is included in said plurality of excitation codebooks, said super codebook comprising bits with a number of bits larger than the number of bits to be transmitted; and selection of the candidate of the excitation codevector from said excitation super codebook is performed corresponding to the candidate of the excitation codevector already selected.

7. A voice-coder for coding speech signals divided into frames spaced with a constant interval, comprising:

   linear prediction analysis means (110) for outputting spectral parameters of input speech signals;
   an adaptive codebook (175) for storing excitation signals determined in the past as adaptive code vectors;
   a plurality of excitation codebooks (180, 190) provided for multi-stage vector quantization of the excitation signal of said input speech signals; and
   a combination searching means for searching a combination of the adaptive codevector and each of the excitation codevectors for every subframe prepared by further division of said frame from said adaptive codebook (175) and each of said excitation codebooks (180, 190), respectively, said combination of the adaptive codevector and each of the excitation codevectors forming an excitation signal of said subframe; characterized in that said combination searching means selects : a first predetermined number (L0) of candidates of adaptive codevectors from said adaptive codebook by using said input speech signal and said spectral parameter;
   a predetermined number (L1, L2) of candidates of excitation codevectors from each of said excitation codebooks, respectively, by using said input speech signal, said spectral parameter and said candidates of said adaptive codevectors; and
   an optimum combination of the adaptive codevector and each of the excitation codevectors forming the excitation signal of said subframe from all the combinations of said candidates of said adaptive codevectors and said candidates of each of said excitation codevectors by using said input speech signal and said spectral parameter, the optimum combination being the one which minimizes a predetermined error function.

8. A voice-coder according to Claim 7, further comprising:

   a gain codebook (220) for quantization of each gain of said adaptive codebook (175) and each of said excitation codebooks (180, 190); wherein
   said input speech signal, said spectral parameter and said gain codebook (220) are used for searching a combination of the adaptive codevector and each of the excitation codevectors which forms the excitation signal of said subframe, from the candidates of said adaptive codevector and said excitation codevectors.

9. A voice-coder according to Claim 7 or 8, wherein:

   at least one or more of excitation super codebooks (390) is included in said plurality of excitation codebooks, said super codebook comprising bits with a number of bits larger than the number of bits to be transmitted; and selection of the candidate of the excitation codevector from said excitation super codebook (390) is performed corresponding to the candidate of the excitation codevector already selected.

10. A voice-coder according to claim 7, wherein said combination searching means comprises:

   subframe division means (120) for generating subframe signals by dividing said frame into subframes;
   first selection means (150) for selecting said candidate of said first predetermined number of said adaptive

codevectors from said adaptive codebook (175) in accordance with said subframe signal and said spectral parameter;

second selection means (160, 170) provided for every said excitation codebook (180, 190) for selecting the candidate of the excitation codevectors of the number predetermined for every excitation codebook, from the corresponding excitation codebook in accordance with said subframe signal, said spectral parameter and the candidate of said adaptive codevector; and

means (200) for searching the candidate of said adaptive codevector and each of said excitation codevectors which forms the excitation signal of said subframe, from the candidate of said adaptive codevector and the candidate of each of said excitation codevectors in accordance with said input speech signal and said spectral parameter.

11. A voice-coder according to Claim 10, wherein:

first and second selecting means (150, 160, 170) select each corresponding candidate, in order, from the candidate with a smaller error; and

said search means (200) searches the candidate of said codevector on the condition of whose error is lowest.

12. A voice-coder according to Claim 10, further comprising:

a gain codebook (200) for quantization of each gain of said adaptive codebook (175) and each of said excitation codebooks (180, 190); wherein

said search means (200) searches the candidate of said codevector by further consulting said gain codebook.

13. A voice-coder according to Claim 11, further comprising:

a gain codebook (220) for quantization of each gain of said adaptive codebook (175) and each of said excitation codebooks (180, 190); wherein

said search means (200) further determines the optimum gain codevector from said gain codebook by consulting said gain codebook (220), and reflects said gain codevector on the adaptive codevector and each of the excitation codevectors which forms said excitation signal.

14. A voice-coder according to any of Claims 10 to 13, wherein, at least one or more of excitation super codebooks (390) is included in said plurality of excitation codebooks, said super codebook comprising bits with a number of bits larger than the number of bits to be transmitted; and

said second selection means (170) corresponding to said excitation super codebook (390) performs selection of the candidate of the excitation codevector from said excitation super codebook according to the candidate of the excitation codevector already selected.

Patentansprüche

1. Verfahren zur Sprachkodierung zum Kodieren von Sprachsignalen, die in Rahmen mit einem konstanten Abstandsintervall unterteilt sind, wobei:

eine adaptive Kodetabelle Anregungssignale speichert, die in der Vergangenheit als adaptive Kodevektoren bestimmt wurden, und mehrere Anregungskodetabellen für eine mehrstufige Vektorquantisierung eines Anregungssignals des Eingangssprachsignals bereitgestellt sind;

ein Spektralparameter des Eingangssprachsignals erhalten wird; und

der Rahmen in Unterrahmen unterteilt wird;

dadurch gekennzeichnet, daß

eine erste feste Anzahl (L_0) von Kandidaten von adaptiven Kodevektoren für jeden Unterrahmen aus der adaptiven Kodetabelle unter Verwendung des Eingangssprachsignals und des Spektralparameters selektiert wird;

eine feste Anzahl (L_1, L_2) von Kandidaten von Anregungskodevektoren für jeden Unterrahmen aus jeder von den Anregungskodetabellen unter Verwendung des Eingangssprachsignals, des Spektralparameters und der

Kandidaten des adaptiven Kodevektors selektiert wird; und
eine optimale Kombination des adaptiven Kodevektors und jedes von den Anregungskodevektoren, die das Anregungssignal des Unterrahmens bilden, aus allen Kombinationen der Kandidaten der adaptiven Kodevektoren und den Kandidaten von jedem der Anregungskodevektoren unter Verwendung des Eingangssprachsignals und des Spektralparameters selektiert wird, wobei die optimale Kombination diejenige ist, welche eine vorbestimmte Fehlerfunktion minimiert.

2. Verfahren zur Sprachkodierung nach Anspruch 1, wobei die Selektion der Kandidaten des adaptiven Kodevektors und jedes von den Anregungskodevektoren jeweils in Reihenfolge von der Selektion des Kandidaten mit einem kleineren Fehler ausgehend durchgeführt wird.

3. Verfahren zur Sprachkodierung nach Anspruch 1 oder 2 wobei:

eine Verstärkungsfaktortabelle zur Durchführung der Quantisierung von Verstärkungsfaktoren der adaptiven Kodetabelle und jeder von den Anregungskodetabellen verwendet wird; und
ein Verstärkungsfaktorkodevektor unter Verwendung der Verstärkungsfaktortabelle bestimmt wird, wenn die Selektion einer Kombination des adaptiven Kodevektors und jedes von den Anregungskodevektoren, die das Anregungssignal des Unterrahmens bilden, aus den Kandidaten des adaptiven Kodevektors und des Anregungskodevektors durchgeführt wird.

4. Verfahren zur Sprachkodierung nach einem der Ansprüche 1 bis 3, wobei:

mindestens eine oder mehrere von Anregungssuperkodetabellen in den mehreren Anregungskodetabellen enthalten sind, die Superkodetabelle Bits mit einer Anzahl größer als die Anzahl zu übertragender Bits aufweist; und
eine Selektion des Kandidaten des Anregungskodevektors aus der Anregungssuperkodetabelle entsprechend dem Kandidaten des bereits selektierten Anregungskodevektor durchgeführt wird.

5. Verfahren zur Sprachkodierung nach Anspruch 1, wobei der Schritt der Selektion der Kombination des adaptiven Kodevektors und jedes von den Anregungskodevektoren, die das Anregungssignal des Unterrahmens bilden, aus den Kandidaten des adaptiven Kodevektors und des Anregungskodevektors ferner die Schritte aufweist:

Ermitteln des optimalen Verstärkungsfaktorkodevektors aus der Verstärkungsfaktortabelle; und
Reflektieren des Verstärkungsfaktorkodevektors auf den adaptiven Kodevektor und jeden von den Anregungskodevektoren, die das Anregungssignal bilden.

6. Verfahren zur Sprachkodierung nach Anspruch 5, wobei:

mindestens eine oder mehrere von den Anregungssuperkodetabellen in den mehreren Anregungskodetabellen enthalten sind, die Superkodetabelle Bits mit einer Anzahl größer als die Anzahl zu übertragender Bits aufweist, und
eine Selektion des Kandidaten des Anregungskodevektors aus der Anregungssuperkodetabelle entsprechend dem Kandidaten des bereits selektierten Anregungskodevektors durchgeführt wird.

7. Sprachkodierer zum Kodieren von Sprachsignalen, die in Rahmen mit einem konstanten Abstandsintervall unterteilt sind, mit:

einer Linearprädiktionsanalyseeinrichtung (110) zum Ausgeben von Spektralparametern der Eingangssprachsignale;
einer adaptiven Kodetabelle (175) zum Speichern von Anregungssignalen, die in der Vergangenheit als adaptive Kodevektoren bestimmt wurden;
mehreren Anregungskodetabellen (180, 190), die für eine mehrstufige Vektorquantisierung des Anregungssignals der Eingangssprachsignale vorgesehen sind; und
einer Kombinationssucheinrichtung zum Suchen einer Kombination des adaptiven Kodevektors und jedes von den Anregungskodevektoren für jeden Unterrahmen, der durch eine weitere Unterteilung des Rahmens erzeugt wird, aus der adaptiven Kodetabelle (175) und aus jeder von den Anregungskodetabellen (180, 190), wobei die Kombination des adaptiven Kodevektors und jedes von den Anregungskodevektoren ein Anregungssignal des Unterrahmens bilden;

dadurch gekennzeichnet, daß die Kombinationssucheinrichtung selektiert:

eine erste vorbestimmte Anzahl (L_0) von Kandidaten von adaptiven Kodevektoren aus der adaptiven Kode-tabelle unter Verwendung des Eingangsprachsignals und des Spektralparameters;
eine vorbestimmte Anzahl (L_1, L_2) von Kandidaten von Anregungskodevektoren aus jeder von den Anregungs-kodetabellen unter Verwendung des Eingangsprachsignals, des Spektralparameters und der Kandidaten von den adaptiven Kodevektoren ; und
eine optimale Kombination des adaptiven Kodevektors und jedes von den Anregungskodevektoren, die das Anregungssignal des Unterrahmens bilden, aus allen Kombinationen der Kandidaten des adaptiven Kode-vektors, und den Kandidaten von jedem der Anregungskodevektoren unter Verwendung des Eingangsprach-signals und des Spektralparameters, wobei die optimale Kombination diejenige ist, welche eine vorbestimmte Fehlerfunüktion minimiert.

8. Sprachkodierer nach Anspruch 7, ferner mit;

einer Verstärkungsfaktorkodetabelle (220) zum Quantisieren jedes Verstärkungsfaktors der adaptiven Kode-tabelle (175) und jeder Anregungskodetabelle (180,190); wobei
das Eingangsprachsignal, der Spektralparameter und die Verstärkungsfaktorkodetabelle (220) zum Suchen einer Kombination des adaptiven Kodevektors und jedes von den Anregungskodevektoren, welche das An-regungssignal des Unterrahmens bilden, aus den Kandidaten des adaptiven Kodevektors und von den Anre-gungskodevektoren verwendet werden.

9. Sprachkodierer nach Anspruch 7 oder 8, wobei

mindestens eine oder mehrere von den Anregungssuperkodetabellen (390) in den mehreren Anregungsko-detabellen enthalten ist, die Superkodetabelle Bits mit einer Anzahl größer als die Anzahl zu übertragender Bits aufweist; und
die Selektion des Kandidaten des Anregungskodevektors aus der Anregungssuperkodetabelle (390) entspre-chend dem Kandidaten des bereits selektierten Anregungskodevektor durchgeführt wird.

10. Sprachkodierer nach Anspruch 7, wobei die Kombinationssucheinrichtung aufweist:

eine Unterrahmenteilungseinrichtung (120) zum Erzeugen von Unterrahmen durch Unterteilen des Rahmens in Unterrahmen;
eine erste Selektionseinrichtung (150) zum Selektieren des Kandidaten der ersten vorbestimmten Anzahl der adaptiven Kodevektoren aus der Anregungskodetabelle (175) in Übereinstimmung mit dem Unterrahmensi-gnal und dem Spektralparameter;
eine für jede Anregungskodetabelle (180, 190) vorgesehene zweite Selektionseinrichtung (160, 170) zum Selektieren des Kandidaten der Anregungskodevektoren mit der für jede Anregungskodetabelle vorbestimm-ten Anzahl, aus der entsprechenden Anregungskodetabelle in Übereinstimmung mit dem Unterrahmensignal, dem Spektralparameter und dem Kandidaten des adaptiven Kodevektors; und
einer Einrichtung (200) zum Suchen des Kandidaten des adaptiven Kodevektors und jedes von den Anre-gungskodevektoren, welcher das Anregungssignal des Unterrahmens bildet, aus dem Kandidaten des adap-tiven Kodevektors und dem Kandidaten jedes von den Anregungskodevektoren in Übereinstimmung mit dem Unterrahmensignal und dem Spektralparameter.

11. Sprachkodierer nach Anspruch 10, wobei:

die ersten und die zweiten Selektionseinrichtungen (150, 160, 170) jeden entsprechenden Kandidaten in einer Reihenfolge ausgehend von dem Kandidaten mit einem kleineren Fehler selektieren; und
die Sucheinrichtung (200) den Kandidaten des Kodevektors unter der Bedingung des kleinsten Fehlers sucht.

12. Sprachkodierer nach Anspruch 10, ferner mit:

einer Verstärkungsfaktorkodetabelle (220) zum Quantisieren jedes Verstärkungsfaktors der adaptiven Kode-tabelle (175) und jeder von den Anregungskodetabellen (180, 190); wobei
die Sucheinrichtung (200) den Kandidaten des Kodevektors durch eine weitere Befragung der Verstärkungs-faktortabelle sucht.

13. Sprachkodierer nach Anspruch 11, ferner mit:

einer Verstärkungsfaktorkodetabelle (220) zum Quantisieren jedes Verstärkungsfaktors der adaptiven Kodetabelle (175) und jeder von den Anregungskodetabellen (180, 190); wobei
die Sucheinrichtung (200) ferner den optimalen Kodevektor aus der Verstärkungsfaktortabelle durch Befragung der Verstärkungsfaktorkodetabelle (220) sucht und den Verstärkungsfaktorkodevektor auf den adaptiven Kodevektor und jeden von den Anregungskodevektoren reflektiert, welcher das Anregungssignal bildet.

14. Sprachkodierer nach einem der Ansprüche 10 bis 13, wobei mindestens eine oder mehrere von den Anregungssuperkodetabellen (330) in den mehreren Anregungstabellen enthalten sind, die Superkodetabelle Bits mit einer Anzahl größer als die Anzahl zu übertragender Bits aufweist, und

die dem Anregungssuperkodetabelle (390) entsprechende zweite Selektionseinrichtung (170) die Selektion des Kandidaten des Anregungskodevektors aus der Anregungssuperkodetabelle entsprechend dem Kandidaten des bereits selektierten Anregungskodevektors durchführt.

Revendications

1. Procédé de codage de la parole pour coder des signaux vocaux divisés en trames espacées d'un intervalle constant, dans lequel :

une table de codage adaptative stockant des signaux d'excitation déterminés par le passé comme étant des vecteurs de code adaptatifs et une pluralité de tables de codage d'excitation pour une quantification de vecteurs à étages multiples d'un signal d'excitation du signal vocal d'entrée sont préparées ;
un paramètre spectral dudit signal vocal d'entrée est obtenu ;
ladite trame est divisée en sous-trames, caractérisé en ce que :
un premier nombre fixe (L0) de candidats aux vecteurs de code adaptatifs est choisi pour chaque dite sous-trame dans ladite table de codage adaptative en utilisant ledit signal vocal d'entrée et ledit paramètre spectral ;
un nombre fixe (L1, L2) de candidats aux vecteurs de code d'excitation est choisi pour chacune desdites sous-trames dans chacune desdites tables de codage d'excitation, respectivement, en utilisant ledit signal vocal d'entrée, ledit paramètre spectral et lesdits candidats auxdits vecteurs de code adaptatifs ; et
une combinaison optimale du vecteur de code adaptatif et de chacun des vecteurs de code d'excitation formant le signal d'excitation de ladite sous-trame est choisie parmi toutes les combinaisons desdits candidats audit vecteur de code adaptatif et lesdits candidats à chacun desdits vecteurs de code d'excitation en utilisant ledit signal vocal d'entrée et ledit paramètre spectral, la combinaison optimale étant celle qui réduit au minimum une fonction d'erreur prédéterminée.

2. Procédé de codage de la parole selon la revendication 1, dans lequel le choix des candidats au vecteur de code adaptatif et chacun des vecteurs de code d'excitation sont effectués, respectivement, dans l'ordre, à partir du choix du candidat présentant le moins d'erreurs.

3. Procédé de codage de la parole selon la revendication 1 ou 2, dans lequel

une table de codage de gain est utilisée pour effectuer la quantification des gains de ladite table de codage adaptative et de chacune desdites table de codage d'excitation, respectivement ; et
un vecteur de code de gain est déterminé en utilisant ladite table de codage de gain lorsque le choix d'une combinaison du vecteur de code adaptatif et de chacun des vecteurs de code d'excitation formant le signal d'excitation de ladite sous-trame parmi les candidats audit vecteur de code adaptatif et audit vecteur de code d'excitation est fait.

4. Procédé de codage de la parole selon l'une quelconque des revendications 1 à 3, dans lequel :

au moins une ou plusieurs super tables de codage d'excitation est ou sont incluse(s) dans ladite pluralité de tables de codage d'excitation, ladite super table de codage comprenant des bits, ce nombre de bits étant supérieur au nombre de bits devant être transmis ; et
le choix du candidat au vecteur de code d'excitation dans ladite super table de codage d'excitation est fait en fonction du candidat au vecteur de code d'excitation déjà choisi.

5. Procédé de codage de la parole selon la revendication 1, dans lequel l'étape de choix de la combinaison du vecteur de code adaptatif et de chacun des vecteurs de code d'excitation formant le signal d'excitation de ladite sous-trame parmi les candidats audit vecteur de code adaptatif et audit vecteur de code d'excitation, comprenant en outre les étapes consistant à :

déterminer le vecteur de code de gain optimal à partir de ladite table de codage de gain ; et
réfléchir ledit vecteur de code de gain sur ledit vecteur de code adaptatif et chacun desdits vecteurs de code d'excitation formant ledit signal d'excitation.

6. Procédé de codage de la parole selon la revendication 5, dans lequel :

au moins une ou plusieurs super tables de codage d'excitation est ou sont incluse(s) dans ladite pluralité de tables de codage d'excitation, ladite super table de codage comprenant des bits, ce nombre de bits étant supérieur au nombre de bits à transmettre ; et
le choix du candidat au vecteur de code d'excitation dans ladite super table de codage d'excitation est fait en onction du candidat au vecteur de code d'excitation déjà choisi.

7. Synthétiseur vocal permettant de coder des signaux vocaux divisés en trames espacées d'un intervalle constant, comprenant :

des moyens d'analyse de prévision linéaire (110) pour sortir des paramètres spectraux des signaux vocaux d'entrée ;
une table de codage adaptative (175) pour stocker des signaux d'excitation déterminés par le passé sous la forme de vecteurs de code adaptatifs ;
une pluralité de tables de codage d'excitation (180, 190) fournie pour la quantification de vecteurs à étages multiples du signal d'excitation desdits signaux vocaux d'entrée ; et
des moyens de recherche de combinaison pour rechercher une combinaison du vecteur de code adaptatif et de chacun des vecteurs de code d'excitation pour chaque sous-trame préparée par une autre division de ladite trame dans ladite table de codage adaptative (175) et chacune desdites table de codage d'excitation (180, 190) respectivement, ladite combinaison du vecteur de code adaptatif et de chacun des vecteurs de code d'excitation formant un signal d'excitation de ladite sous-trame ; caractérisé en ce que lesdits moyens de recherche de combinaison choisissent :
un premier nombre (L0) prédéterminé de candidats aux vecteurs de code adaptatifs dans ladite table de codage adaptative en utilisant ledit signal vocal d'entrée et ledit paramètre spectral ;
un nombre prédéterminé (L1, L2) de candidats aux vecteurs de code d'excitation dans chacune desdites tables de codage d'excitation, respectivement, en utilisant ledit signal vocal d'entrée, ledit paramètre spectral et lesdits candidats auxdits vecteurs de code adaptatifs ; et
une combinaison optimale du vecteur de code adaptatif et de chacun des vecteurs de code d'excitation formant le signal d'excitation de ladite sous-trame parmi toutes les combinaisons desdits candidats audit vecteur adaptatif et desdits candidats à chacun desdits vecteurs de code d'excitation en utilisant ledit signal vocal d'entrée et ledit paramètre spectral, la combinaison optimale étant celle qui réduit au minimum une fonction d'erreur prédéterminée.

8. Synthétiseur vocal selon la revendication 7, comprenant en outre :

une table de codage de gain (220) pour quantifier chaque gain de ladite table de codage adaptative (175) et chacune desdites tables de codage d'excitation (180, 190) ; dans lequel
ledit signal vocal d'entrée, ledit paramètre spectral et ladite table de codage de gain (220) sont utilisés pour rechercher une combinaison du vecteur de code adaptatif et de chacun des vecteurs de code d'excitation qui forme le signal d'excitation de ladite sous-trame, parmi les candidats audit vecteur de code adaptatif et auxdits vecteurs de code d'excitation.

9. Synthétiseur vocal selon la revendication 7 ou 8, dans lequel :

au moins une ou plusieurs super tables de codage d'excitation (390) est ou sont incluse(s) dans ladite pluralité de tables de codage d'excitation, ladite super table de codage comprenant des bits, le nombre de ces bits étant supérieur au nombre de bits à transmettre ; et
le choix du candidat au vecteur de code d'excitation dans ladite super table de codage d'excitation (390) est

fait en fonction du candidat au vecteur de code d'excitation déjà choisi.

10. Synthétiseur vocal selon la revendication 7, dans lequel lesdits moyens de recherche de combinaisons comprennent :

des moyens (120) de division de sous-trames pour générer des signaux de sous-trames en divisant ladite trame en sous-trames ;

des premiers moyens de sélection (150) pour choisir ledit candidat parmi ledit premier nombre prédéterminé de candidats auxdits vecteurs de code adaptatifs dans ladite table de codage adaptative (175) en correspondance avec ledit signal de sous-trame et ledit paramètre spectral ;

des deuxièmes moyens de sélection (160, 170) fournis pour chaque table de codage d'excitation (180, 190) pour choisir le candidat aux vecteurs de code d'excitation parmi le nombre prédéterminé pour chaque table de codage d'excitation, dans la table correspondante de codage d'excitation en correspondance avec ledit signal de sous-trame, ledit paramètre spectral et le candidat audit vecteur de code adaptatif ; et

des moyens (200) pour rechercher le candidat audit vecteur de code adaptatif et à chacun desdits vecteurs de code d'excitation qui forme le signal d'excitation de ladite sous-trame, parmi le candidat audit vecteur de code adaptatif et au candidat à chacun desdits vecteurs de code d'excitation en correspondance avec ledit signal vocal d'entrée et ledit paramètre spectral.

11. Synthétiseur vocal selon la revendication 10, dans lequel :

des premier et deuxième moyens de sélection (150, 160, 170) choisissent chaque candidat correspondant, dans l'ordre, sous condition d'erreur minimum ; et

lesdits moyens de recherche (200) recherchent le candidat audit vecteur de code sous condition d'erreur minimum.

12. Synthétiseur vocal selon la revendication 10, comprenant en outre :

une table de codage adaptative de gain (220) pour quantifier chaque gain de ladite table de codage adaptative (175) et de chacune desdites table de codage d'excitation (180, 190) ; dans lequel

lesdits moyens de recherche (200) recherchent le candidat audit vecteur de code en consultant en outre ladite table de codage de gain.

13. Synthétiseur vocal selon la revendication 11, comprenant en outre :

une table de codage adaptative de gain (220) pour quantifier chaque gain de ladite table de codage adaptative (175) et de chacune desdites table de codage d'excitation (180, 190) ; dans lequel

lesdits moyens de recherche (200) déterminent en outre le vecteur de code de gain optimal à partir de ladite table de codage de gain en consultant ladite table de codage de gain (220), et réfléchit ledit vecteur de code de gain sur le vecteur de code adaptatif et chacun des vecteurs de code d'excitation formant ledit signal d'excitation.

14. Synthétiseur vocal selon l'une quelconque des revendications 10 à 13, dans lequel au moins une ou plusieurs super tables de codage d'excitation (390) est ou sont incluse(s) dans ladite pluralité de tables de codage adaptative d'excitation, ladite super table de codage comprenant des bits, le nombre de ces bits étant supérieur au nombre de bits à transmettre ; et

ledit deuxième moyen de sélection (170) correspondant à ladite super table de codage d'excitation (390) choisit le candidat au vecteur de code d'excitation dans ladite super table de codage d'excitation en fonction du candidat au vecteur de code d'excitation déjà choisi.
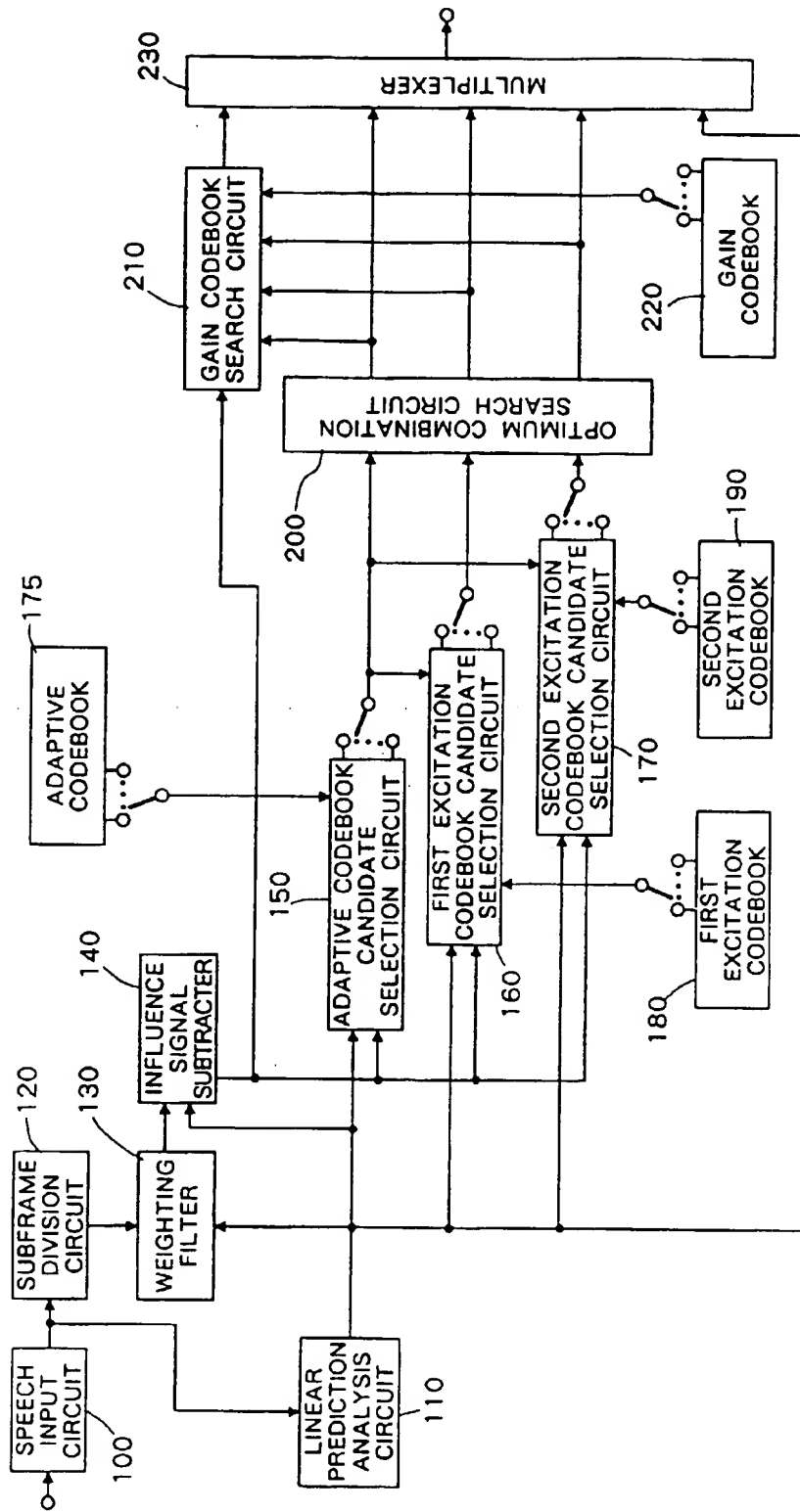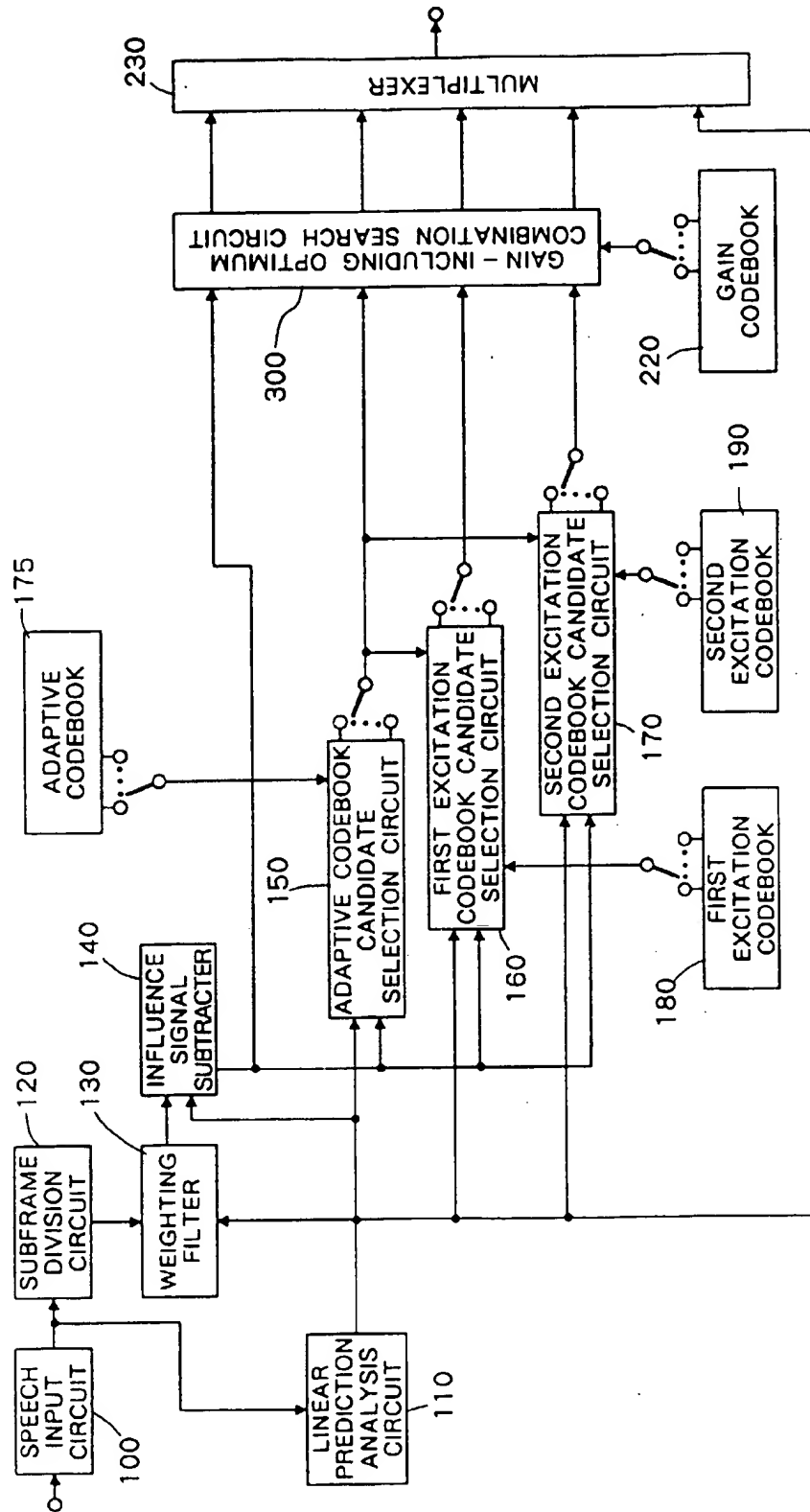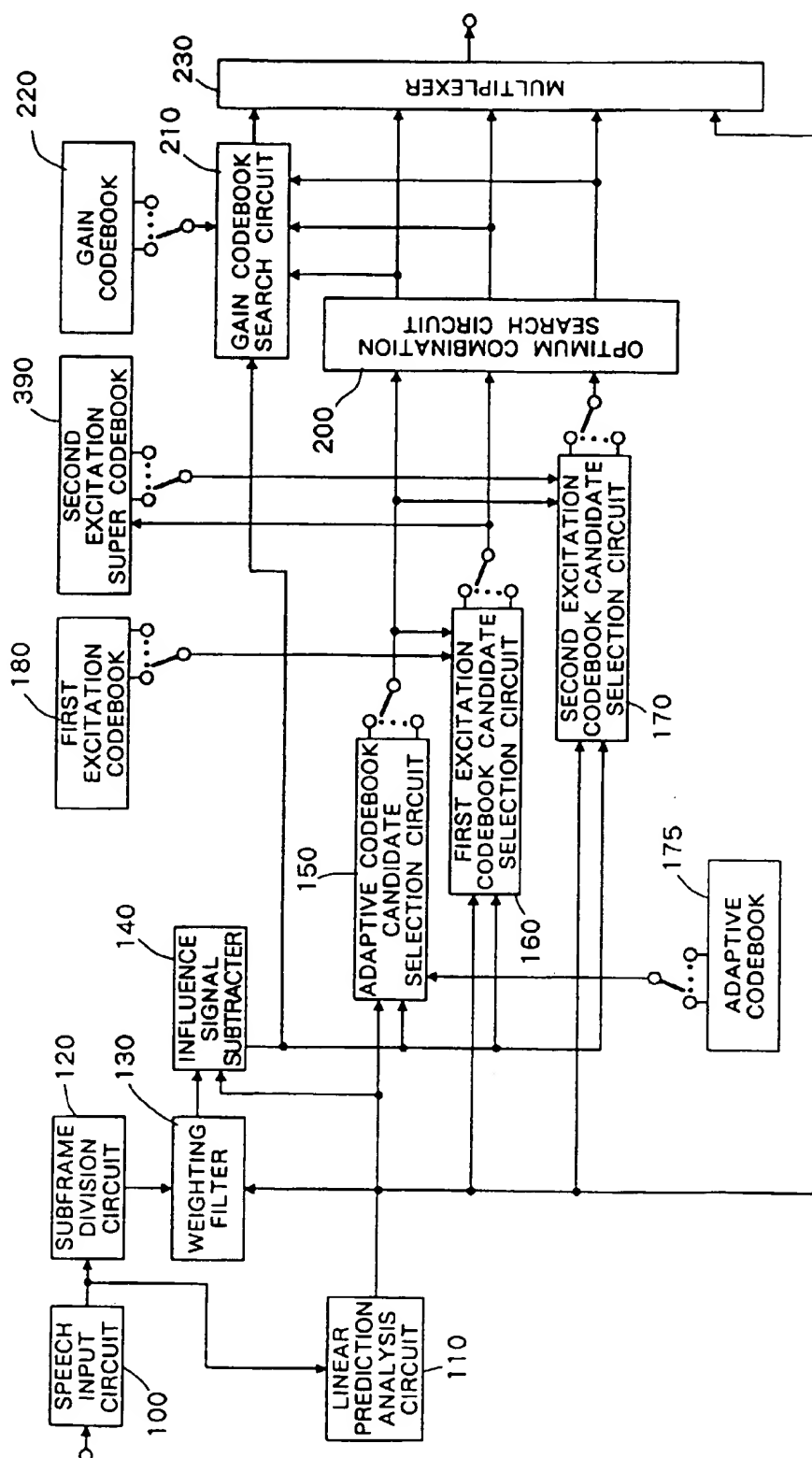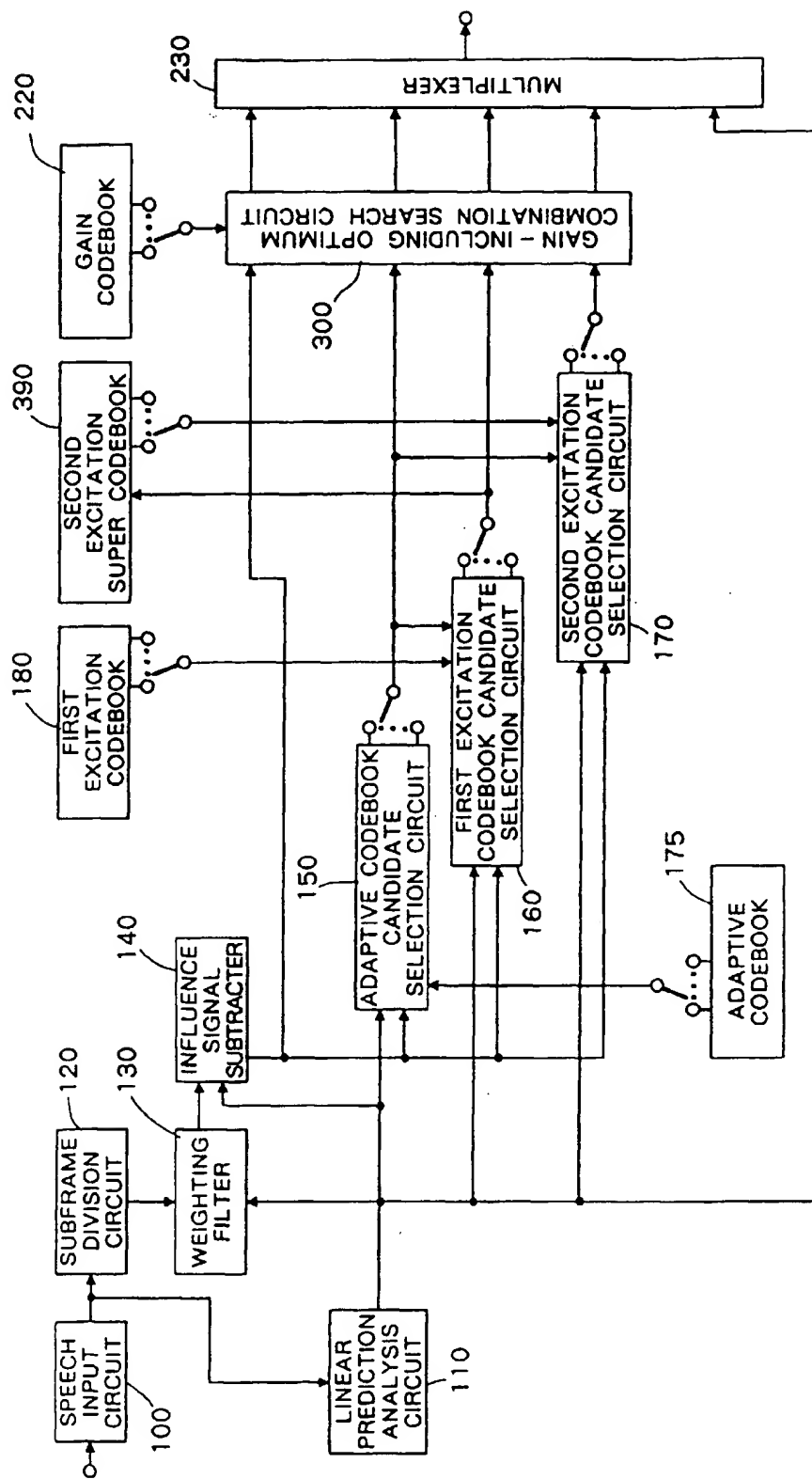
FIG. 1

FIG. 2

FIG. 3

FIG. 4